BMEGUI3.0.0
User Manual

BMEGUI3.0.0

Last Edited on: 01/20/2012

BMElab

Dept. of Environmental Sciences and Engineering

School of Public Health

University of North Carolina

Contents

Contents 2
1 Introduction 6
1.1 About BMEGUI 6
1.2 Download and Installation 6
1.3 Software Requirement 6
2 Starting up BMEGUI 7
3 Data Preparation 7
3.1 Workspace and Data File 7
3.1.1 Workspace 7
3.1.2 Data File 7
3.1.3 River Network File 8
3.2 Data Format 8
3.2.1 GeoEAS Format 8
3.2.2 CSV Format 8
3.2.3 CSV Format without Header Line for River Network 8
3.2.4 Shape File Format for River Network 9
3.3 Required Data Fields 9
3.4 Station ID and System ID 9
3.5 Data File Example 10
3.5.1 GeoEAS Format 10
3.5.2 CSV Format 10
3.6 Hard Data and Soft Data 10
3.6.1 Example (CSV Format) of hard and soft data 12
3.6.2 Example (CSV Format) of river network data 12
4 Getting Started with BMEGUI 13
4.1 Input File and Directory Dialog Box 13
4.2 Input River Network File (River Metric) 14
4.3 Dialog Box 1 (Data Field) 15
4.3.1 Basic Operation 15
4.3.2 Data File with Soft data 16
4.4 Dialog Box 2 (Data Distribution) 17
4.4.1 Basic Operation 17
4.4.2 Data Transformation Method 19
4.4.3 Log of Zero and Negative Value Setting 19
4.4.4 Soft Data in Histogram 20
4.5 Dialog Box 3 (Exploratory Data Analysis) 21
4.5.1 Basic Operation 21
4.5.2 Data Aggregation 23
4.5.3 Create Point Layer File 25
4.6 Dialog Box 4 (Mean Trend Analysis) 26
4.6.1 Basic Operation 26
4.6.2 Calculate Mean Trend Using User-defined Parameters 30
4.6.3 Create Point Layer File 31
4.7 Dialog Box 5 (Space/Time Covariance Analysis) 31

4.7.1 Basic Operation 31
4.7.2 Calculate Experimental Covariance 34
4.7.3 Covariance Model 36
4.7.4 Automatic Covariance Model fitting 38
4.8 Dialog Box 6 39
4.8.1 Basic Operation 39
4.8.2 BME Parameters 41
4.8.3 Estimation Parameters (Spatial Distribution) 42
4.8.4 BME Spatial Estimation 43
4.8.5 BME Mask, Map, Grid, and Color Setting 44
4.8.6 BMEGUI Home Page and Help 45
4.8.7 Create Output Files (Point Layer File and ArcASCII File) 47
4.8.8 Estimation Parameters (Temporal Distribution) 48
4.8.9 BME Temporal Estimation 49
4.8.10 Show, Close, and Delete Maps (or Time Series Plots) 51
4.8.11 Hide and Display Failed Estimation Point 52
4.9 Quitting from BMEGUI 54
5 Interaction with ArcGIS 55
5.1 Details of ArcGIS Files 55
5.2 Coordinate System of ArcGIS Files 55
6 Advanced Topics 57
6.1 Data Error Handling 57
6.2 BMEGUI Parameter File and Estimation Files 58
7 Troubleshooting errors 60
7.1 Data Error file due to an inappropriate new line character 60

List of Figures

Figure 1: River network csv file and network plot 13
Figure 2: Input file (s) and directory dialog box. 13
Figure 3: (A) Euclidian distances and (b) river distances 14
Figure 4: Dialog Box 1 (Data Field) 16
Figure 5: Dialog Box 1 (Data Field) - To use the soft data, check the "Use Datatype" check box 17
Figure 6: Dialog Box 2 (Data Distribution) 18
Figure 7: Use Log-transformed data 19
Figure 8: Settings for the log of negative and zero data values 20
Figure 9: Dialog Box 3 (Exploratory Data Analysis) 22
Figure 10: "Spatial Distribution" tab - Methods to select specific times 23
Figure 11: Example of data aggregation with 10 time-unit aggregation period. (1) raw data and (2) aggregated data 24
Figure 12: Data aggregation 25
Figure 13: The "Create Point Layer" button and the message box 26
Figure 14: Dialog Box 4 (Mean Trend Analysis) 27
Figure 15: Calculating the global mean trend and removing it from the data 29
Figure 16: The mean trend smoothing parameters and the "Recalculate Mean Trend" button 31
Figure 17: Dialog Box 5 (Space/Time Covariance Analysis) 33
Figure 18: Calculating experimental covariance by modifying the number of the lags 34
Figure 19: Calculating experimental covariance values by directly entering the lags and the lag tolerances 36
Figure 20: Covariance model parameter settings 38
Figure 21: Automatic covariance model fitting 39
Figure 22: Dialog Box 6 (BME Estimation) 40
Figure 23: BME Parameters 42
Figure 24: Estimation parameters for the BME spatial estimation 43
Figure 25: List of BME estimation maps 44
Figure 26: Maps of BME mean estimates and BME error variances 46
Figure 27: BMEGUI box 6 showing BMEGUI generated maps and buttons for creating ArcGIS compatible and vector data file using BMEGUI output. 47
Figure 28: Enter name of the file to save as BMEGUI output (a), message after creating ArcGIS compatible (ArcASCII) file (b), and point layer (csv) files (c) 48
Figure 29: Estimation and Display Parameters used for the BME temporal estimation 49
Figure 30: List of estimated time series 50
Figure 31: The time series plot at a specific monitoring location 51
Figure 32: The "Close Tab", "Show", and "Delete" buttons and the message box to confirm the deletion. 52
Figure 33: Message showing the number of failed estimation points. 53
Figure 34: Failed Estimation Points (black dots) on the estimation map 54
Figure 35: The message dialog box to confirm whether to quit BMEGUI. 54
Figure 36: ArcGIS warning message 56
Figure 37: The various message dialog boxes that display when data errors are detected. 58

Figure 38: Error message due to an inappropriate new line character..... 60
Figure 39: ConTEXT editor..... 61

1 Introduction

1.1 About BMEGUI

BMEGUI is the software providing a Graphical Users Interface (GUI) to the Bayesian Maximum Entropy (BME) advanced functions of Space/Time geostatistical analysis. Using this software, the user has access to an easy-to-use interface for the analysis of space/time data.

BMEGUI version 3.0.0 uses BMElib 2.0b, and python 2.5.

1.2 Download and Installation

To install BMEGUI, go to the BMEGUI website at: http://www.unc.edu/depts/case/BMEGUI/ and select version 3.0.0 from the list, which gives you access to the installation package and installation manual. Follow the instructions on the installation manual.

1.3 Software Requirement

BMEGUI uses the following software modules. Before using the software you need to install all software modules.

- GTK 2.10.11
- FreeType
- Python Libraries
  - PyCairo
  - PyGObject
  - PyGTK
  - NumPy
  - SciPy
  - Matplotlib
  - Pywin
- MATLAB Component Runtime (MCR)

2 Starting up BMEGUI

BMEGUI can be started by clicking on the BMEGUI shortcut on desktop. You can conduct the basic exploratory data analysis and Space/Time geostatistical estimation for air, water or any other media for any environmental agent i.e. arsenic in water, SO_2 in air etc. You can create ArcASCII or Comma-Separated Value (CSV) output data files for further processing BMEGUI output in ArcGIS, MATLAB and/or any other compactible software.

3 Data Preparation

3.1 Workspace and Data File

In order to use BMEGUI, you need to specify two input parameters, "Workspace" and "Data File". Workspace is a directory which is used to store all the files BMEGUI creates during the analysis. Data File is a file containing the space/time data available, including the measurement values, their space/time coordinates, and information on measurement errors. You need another data file in case you are analyzing river metric.

3.1.1 Workspace

Workspace is used to store all the files BMEGUI creates during the analysis. The followings are the list of the files stored in Workspace.

- BMEGUI parameter files (.ysp)
- BMEGUI estimation files (.yme, .yse)
- Initial parameter files (.py, .pyc)

During the analysis, all the estimation parameters and results are stored in the Workspace. If the user quits the BMEGUI and executes it again using the same Workspace and Data File, then all the estimation parameter settings and results that were saved are automatically used. If the user modifies the estimation parameters during the second analysis, then all the stored parameters and results obtained in the first analysis are erased and overwritten for the current analysis. When that happens, the BMEGUI pops up a dialog box to ask the user if they would like to overwrite the earlier results or not.

3.1.2 Data File

Data File is a file containing the space/time data available, including the measurement values, their space/time coordinates, and information on measurement errors. Currently, BMEGUI supports following two data formats.

- GeoEAS format
- Comma-Separated Value (CSV) format **with header**

The GeoEAS format is the default file format for BMElib packages. BMElib users are able to use the data file prepared for BMElib without any modification.

3.1.3 River Network File

A river network file is a file containing the x-y (e.g. longitude and latitude) coordinates of points defining river reaches. Currently, BMEGUI supports river network files that have one of the following two data formats.

- ArcGIS shape file (*.shp) format
- Coma Separated Value (CSV) format **without header**

  A river network file in *.csv format contains two columns providing the X and Y (e.g. longitude and latitude) coordinates of points delineating river reaches. Each river reach is separated from the next river reach by a line containing two 'NaN' values. A line with two 'NaN' values is added at the end of the last reach, followed by a last line containing the X and Y coordinate of the river outlet. The river outlet coordinate must correspond to the coordinate of the downstream most point of the river network.

3.2 Data Format

As explained in 3.1.2, GeoEAS format and CSV format are supported in BMEGUI. The details of each data format are listed below.

3.2.1 GeoEAS Format

GeoEAS format data must be prepared in the following manner.

- 1st line: File description
- 2nd line: Number of data column
- 3rd line to (3 + number in 2nd line) line: Name of data column
- Tab separated data
- File extension: .txt

3.2.2 CSV Format

CSV format data must be prepared in the following manner.

- 1st line: Comma separated data column name
- Comma separated data
- File extension: .csv

3.2.3 CSV Format without Header Line for River Network

CSV format data for river network must be prepared in the following manner.

- ❖ No data column name (no header)
- ❖ Comma separated data
- ❖ Last line should have river/stream outlet coordinates
- ❖ File extension: .csv

3.2.4 Shape File Format for River Network

A shapefile stores nontopological geometry and attribute information for the spatial features in a data set. The geometry for a feature is stored as a shape comprising a set of vector coordinates.

3.3 Required Data Fields

Since BMEGUI deals with space/time data, the Data File must have at least four data columns; namely the X field, Y field, T field, and data value field. The X field and Y field are used to specify the spatial coordinate. Currently BMEGUI supports only two-dimensional spatial coordinates. The T field is used to describe the time when the measurement are taken. The Data value field corresponds to actual measurement values.

- ❖ X field, Y field: Spatial Coordinates (i.e. longitude, latitude)
- ❖ T field: Time when the measurement are taken
- ❖ Data value field: Measurement values

If the data is purely spatial (i.e. no changes over time), then the user still needs to prepare the T field using a fixed arbitrary value (i.e. indicating that all values were collected at the same time). Conversely, if data is purely temporal (i.e. a time series), then the user still needs to prepare the X field and Y field using some fixed arbitrary values (i.e. indicating that all values were collected at the same spatial location).

In case you want to analyze water quality parameters in river/stream(s) then you have to prepare another data file for river network coordinates. BMEGUI will ask to provide such file if you click on 'Use River Network' check button.

3.4 Station ID and System ID

In addition to the required data fields described in 3.3, the user may want to use a user-defined station ID for each monitoring location (site). The station ID is a unique identification alphanumeric string that is used to identify monitoring locations in various plots of the BMEGUI as well as in its drop-down lists in the third and sixth dialog boxes. Alphanumeric values (0-9, a-z and A-Z) can be used for station ID. To enter user-defined station IDs, the user has to prepare an additional station ID column in the Data File. If the Data File does not have a station ID column, then BMEGUI creates the system ID.

The system ID is automatically assigned to each monitoring location in order to help the user select one specific monitoring location from the lists in the third and sixth dialog boxes. The system ID is a sequential number starting from one.

3.5 Data File Example

3.5.1 GeoEAS Format

```
Tetrachloroethene (micrograms per liter) in New Jersey
7
LONGITUDE
LATITUDE
NUMDAYS
YEAR
DATATYPE
VAL1
VAL2
-74.5278  40.5594  880  2001  0  0.01  0.01
-74.7781  40.2217  376  2000  0  0.01  0.01
```

3.5.2 CSV Format

```
LONGITUDE, LATITUDE,NUMDAYS, YEAR,DATATYPE,VAL1,VAL2
-74.5278,40.5594,880,2001,0,0.01,0.01
-74.7781,40.2217,376,2000,0,0.01,0.01
```

3.6 Hard Data and Soft Data

Hard data correspond to measurements without errors (or with errors that are small enough to be ignored). Soft data correspond to measurements with an associated uncertainty (for example data with appreciable measurement errors). The uncertainty associated with soft data is described by means of a statistical distribution (for example uniform, Gaussian, etc.).

BMEGUI supports the following five data types.

- Hard data
- Soft data with uniform distribution
- Soft data with Gaussian distribution
- Soft data with Triangular distribution
- Soft data with truncated Gaussian distribution

When using the default settings, BMEGUI assumes that the data file only contains hard data, and in that case it uses only the fields described so far (i.e. the X field, the Y field, the T field, the optional ID field, and the Data field containing the hard data values.

However, when using a combination of hard and soft data, then BMEGUI requires that the Data field be replaced by the following five fields: The Data type field, the Value1 field, the Value2 field, the Value3 field, and the Value4 field. The Data type field is used to specify the type of data. The Value1, Value2, Value3, and Value4 fields are used to describe the data, as follow:

- Hard data
    - Data Type: 0
    - Value1 Field: The true value (e.g. a measurement without error)
    - Value2 Field: Same as Value 1
    - Value3 Field: Same as Value 1 or dummy values
    - Value4 Field: Same as Value 1 or dummy values
- Soft uniform data
    - Data Type: 1
    - Value1 Field: Lower bound of the interval for the true value
    - Value2 Field: Upper bound of the interval for the true value
    - Value3 Field: Lower bound of the interval for the true value or dummy values
    - Value4 Field: Upper bound of the interval for the true value or dummy values
    -
- Soft Gaussian data
    - Data Type: 2
    - Value1 Field: Mean (also called expectation) of the true value.
    - Value2 Field: Standard deviation of the true value around its mean
    - Value3 Field: Mean of the true value or dummy values
    - Value4 Field: Standard deviation of the true value around its mean or dummy values
- Soft Triangular data
    - Data Type: 3
    - Value1 Field: Lower limit.
    - Value2 Field: Upper limit
    - Value3 Field: Mode
    - Value4 Field: Lower limit or dummy values
- Soft Truncated Gaussian data
    - Data Type: 4

- Value1 Field: Mean (also called expectation) of the true value.
- Value2 Field: Standard deviation of the true value around its mean
- Value3 Field: Lower truncation point value
- Value4 Field: Upper truncation point value

3.6.1 Example (CSV Format) of hard and soft data

Table 1: Hard and soft data in csv format

X	Y	Time	SiteID	DataType	Val1	Val2	Val3	Val4
-74.1101	39.93762	1	22	0	37.3	37.3	37.3	37.3
-74.0907	39.91373	1	21	0	73.7	73.7	73.7	73.7
-74.1293	39.81845	40	12	1	24.5	51.45	24.5	51.5
-74.1018	39.80984	40	9	1	21.9	45.99	21.9	46
-74.1532	39.79262	40	4	2	18.1	38.01	18.1	38
-74.1604	39.82512	40	13	2	18.1	38.01	18.1	38
-74.0551	40.04123	43	34	2	33	69.3	33	69.3
-74.1263	39.88206	188	18	3	3.4	7.14	5.27	7.14
-74.1413	39.84012	188	15	3	3.4	7.14	5.27	7.14
-74.1101	39.93762	220	22	4	30.8	64.68	47.7	60.7
-74.114	39.99956	250	29	4	9.3	19.53	14.4	16.5
-74.056	40.03956	270	33	4	4.9	10.29	7.6	8.29

If using hard and soft data then we have to have four data columns i.e. val1, val2, val3, and val4 However, BMEGUI uses only yellow colored columns. It is clear that in case of 0, 1, 2 data types, BMEGUI uses only Val1, and Val2 column so user can put any garbage values in Val3 and Val4 columns. In case of data type 3 BMEGUI uses only 3 columns so fourth column can have any garbage value. However, in case of data type equal 4; all columns are used in space-time estimation.

3.6.2 Example (CSV Format) of river network data

The CSV file shown below corresponds to a river network consisting of three river reaches. Each river reach is delineated with three points located along that river reach. The CSV file contains two columns listing the X, Y coordinates of the points delineated each river reach. Each river reach is ended by a line with NaN values. The last line has the X-Y coordinates of the river network outlet.

X	Y
1	7
2	5
3	3
NaN	NaN
5	7
4	5
3	3
NaN	NaN
3	3
3	2
3	1
NaN	NaN
3	1

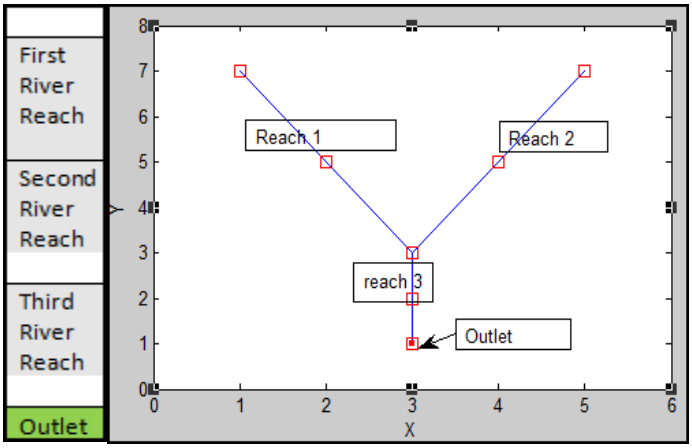


Figure 1: River network csv file and network plot

4 Getting Started with BMEGUI

4.1 Input File and Directory Dialog Box

BMEGUI can be started by clicking on the BMEGUI shortcut on desktop. BMEGUI pops up a window as shown below in figure 1.

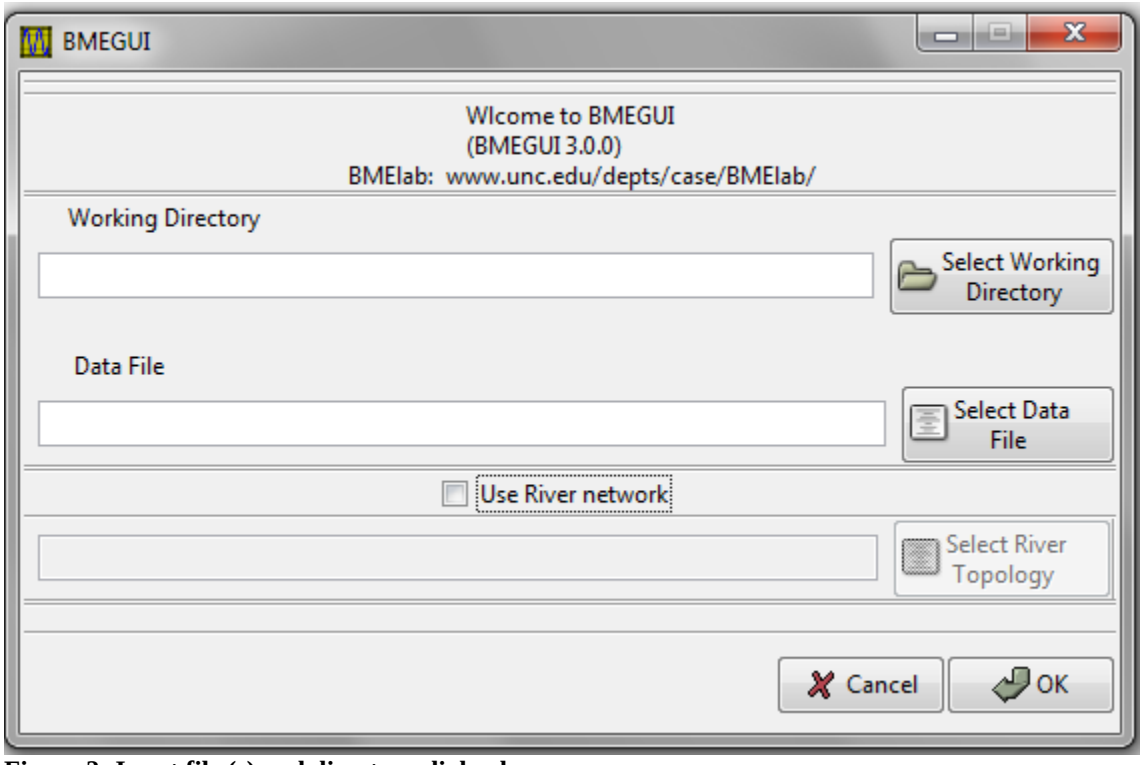


Figure 2: Input file (s) and directory dialog box.

You can provide the working directory and the data file to BMEGUI by clicking on the 'Select Working Directory' and 'Select Data File' buttons. Leave the 'Use River Network' unchecked in order to use the Euclidean metric. This is the default settings, which calculates the straight line distance between points.

4.2 Input River Network File (River Metric)

Most geostatistical studies use the Euclidean, or 'across land', metric to calculate the distance between points. However, the choice of the distance metric is critical, because the distance metric affects correlation functions (such as the covariance), as well as the selection of the estimation neighborhood. Recent studies have shown that when dealing with water quality parameters measured along a river, it may be better to use the river metric. This will result in BME maps that are more accurate and realistic than those obtained using an Euclidean distance.

Using the river metric, the distance between two points is calculated by finding the shortest path between these points along the river, as opposed to a straight line distance. Hence the river metric leads to distances that are always greater than that obtained with the Euclidean metric, with an increase in distance that captures the topology of the river network.

(A) Euclidian distances and (b) river distances

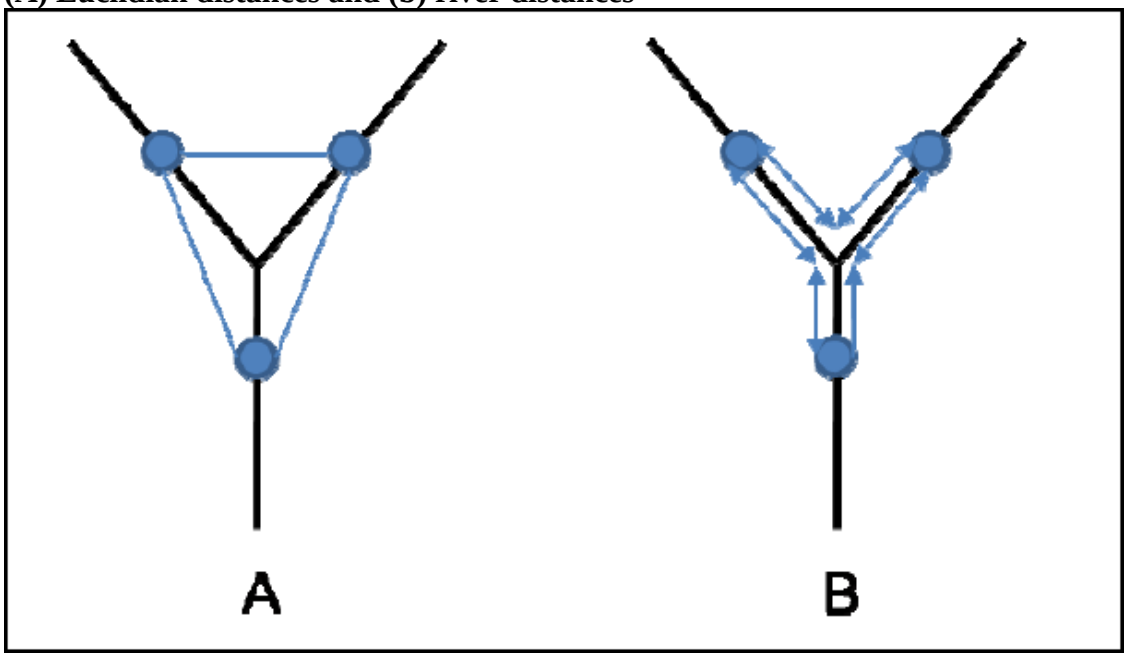


Figure 3: (A) Euclidian distances and (b) river distances

Note that Euclidean distance and isotropic river distance both meet the qualifications of a metric, therefore the term 'metric' and 'distance' are used interchangeably in this document. The term 'metric' is used to describe a distance that meets the following criteria for spatial points ***s***, ***s'***, and ***s"***

$d(\boldsymbol{s}, \boldsymbol{s'}) \geq 0$ (non-negativity)
$d(\boldsymbol{s}, \boldsymbol{s'}) = 0$ if and only if $\boldsymbol{s} = \boldsymbol{s'}$ (identity)
$d(\boldsymbol{s}, \boldsymbol{s'}) = d(\boldsymbol{s'}, \boldsymbol{s})$ (symmetry)

$d(\boldsymbol{s}, \boldsymbol{s'}) \leq d(\boldsymbol{s}, \boldsymbol{s''}) + d(\boldsymbol{s''}, \boldsymbol{s'})$ (triangle inequality)

In order to use the river metric, the user just need to check the 'Use River Network' button, and provide a file that describes the topology of the river network. BMEGUI will then automatically use covariance models that are permissible for the river metric, so that the analysis is mathematically sound.

The topology of a river network is represented by a directed tree of river reaches with zero width. Such representation is adequate only for downstream combining stream networks with somewhat narrow reaches but may not be adequate for wider water bodies such as connected estuaries or lakes. In order to construct this river topology, BMEGUI needs a river network data file that delineates individual river reaches of the river network, as well as the point corresponding to the outlet of this river network. The river network must be such that there is single river reach at the downstream end of the river network, and the outlet provided by the user must be the downstream end of that reach. BMEGUI will then attempt to construct the river topology by starting from the outlet, and finding each river reach that is connected to that outlet. Every time BMEGUI encounters a point where the river divides into two confluent reaches, BMEGUI will designate that point as a confluence point, and will delineate the two upstream reaches up to their upstream confluence point. The process is repeated up until all the upstream reaches have been identified. When this process is finished BMEGUI will have reorganized the set of river reaches into a directed tree of downstream only combining reaches.

In practice, there may be many problems associated with river networks. BMEGUI tests for the following river network problems:

1. Network outlet error:
   (a) Network outlet not connected with the river network
   (b) Network outlet located at the end of multiple river reaches
2. Broken river network
   Missing links between river reaches leads to river reaches that are not connected to the outlet, or simply put, to a broken river network.

There may be many other types of problems in the river network that cannot be handled by BMEGUI. The user is therefore advised to plot and visualize the river network in any GIS software and make any necessary correction needed to eliminate broken links in the river network. See tutorial 7 for an example of a geostatistical analysis using a river network.

4.3 Dialog Box 1 (Data Field)

4.3.1 Basic Operation

Dialog Box 1 (Data Field) shown in **Error! Reference source not found.** is used to elect which data columns of the data file will be used in the analysis, and to enter the units of these data columns, as well as the name of parameter being mapped.

The “Working Directory/Data File” section shows the directories of Workspace and Data File used in the analysis, so that the user can verify these directories.

In the “Data Field Setting” section, the user can select which data columns of the data file are used in the analysis. As explained in 3.3, the data file must have at least four data columns corresponding the following fields:

- X field, Y field
- T field
- Data value field

The user can select the name of the data column for the X Field, Y Field, T Field, and Data Field using the corresponding drop down menus. In addition, the user can select the data column for the station ID (ID field). The default setting of the ID field is “Automatic ID”, which automatically assigns sequential ID to each measurement locations. If the data file does not have a column specifying user-defined IDs, then use the default setting.

In the “Unit/Name” section, the user can directly input the unit for the spatial coordinate, the time event, and the measurement values, as well as the name of the parameter being mapped. The units and the name of the parameter being mapped are only used in the labels of the plots generated by the BMEGUI.

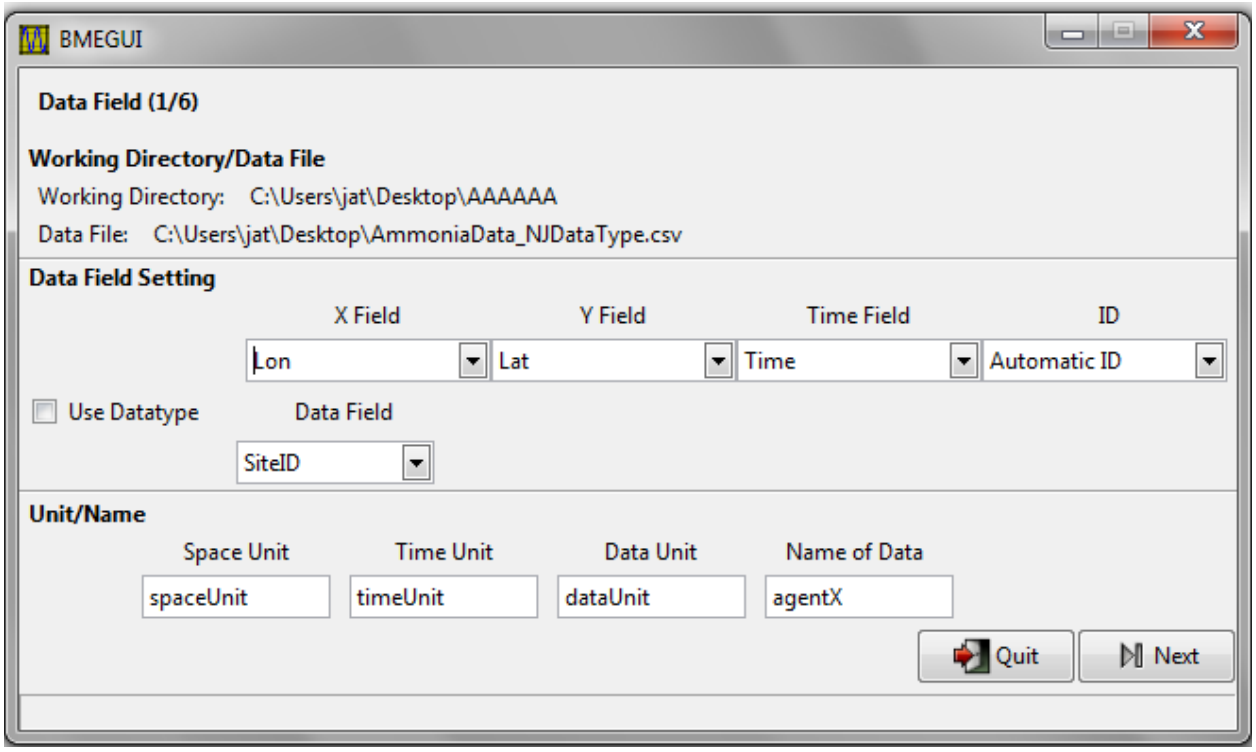


Figure 4: Dialog Box 1 (Data Field)

4.3.2 Data File with Soft data

As explained in 0, BMEGUI supports space/time analysis using soft data in addition to hard data. To use soft data, the user needs to specify which columns of the data file

correspond to the data type field, the value1 field, the value2 field, the value3 field, and the value4 field. The procedure is as follow:

1) Check the "Use Datatype" check box, then drop down boxes for "Data Type", "Value1 Field", "Value2 Field", "Value3 Field", and "Value4 Field" will appear (**Error! Reference source not found.**)
2) Select the appropriate data columns for "Data Type", "Value1 Field", "Value2 Field", "Value3 Field", and "Value4 Field"
3) Click "Next" to move to the second dialog box.

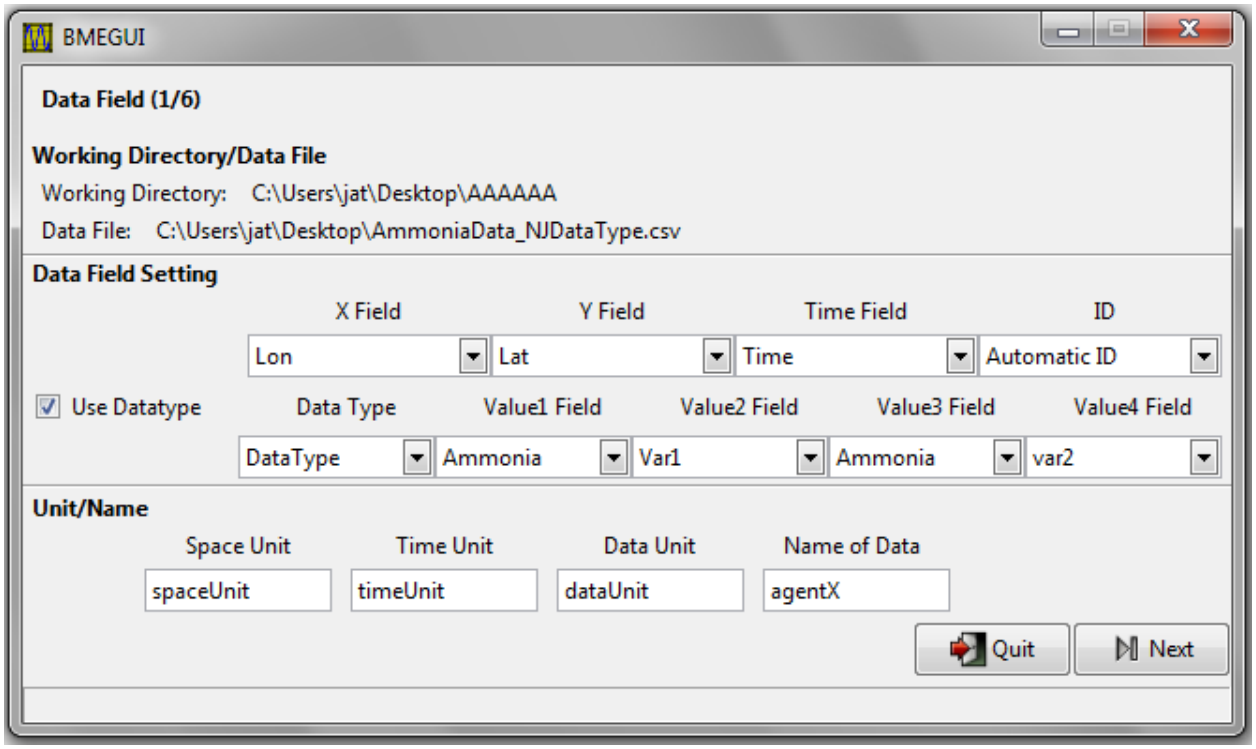


Figure 5: Dialog Box 1 (Data Field) - To use the soft data, check the "Use Datatype" check box

4.4 Dialog Box 2 (Data Distribution)

4.4.1 Basic Operation

Dialog Box 2 (Data Distribution) shown in **Error! Reference source not found.** is used o check the statistical distribution of the data.

The "Statistics" section displays the basic statistics of the raw data and of the log-transformed data.

The "Histogram" section displays the histogram of the raw and log-transformed data. By switching the tabs between "Raw data" and "Log Data", the user can switch histograms. The user can also modify the settings for the log of negative and zero data values.

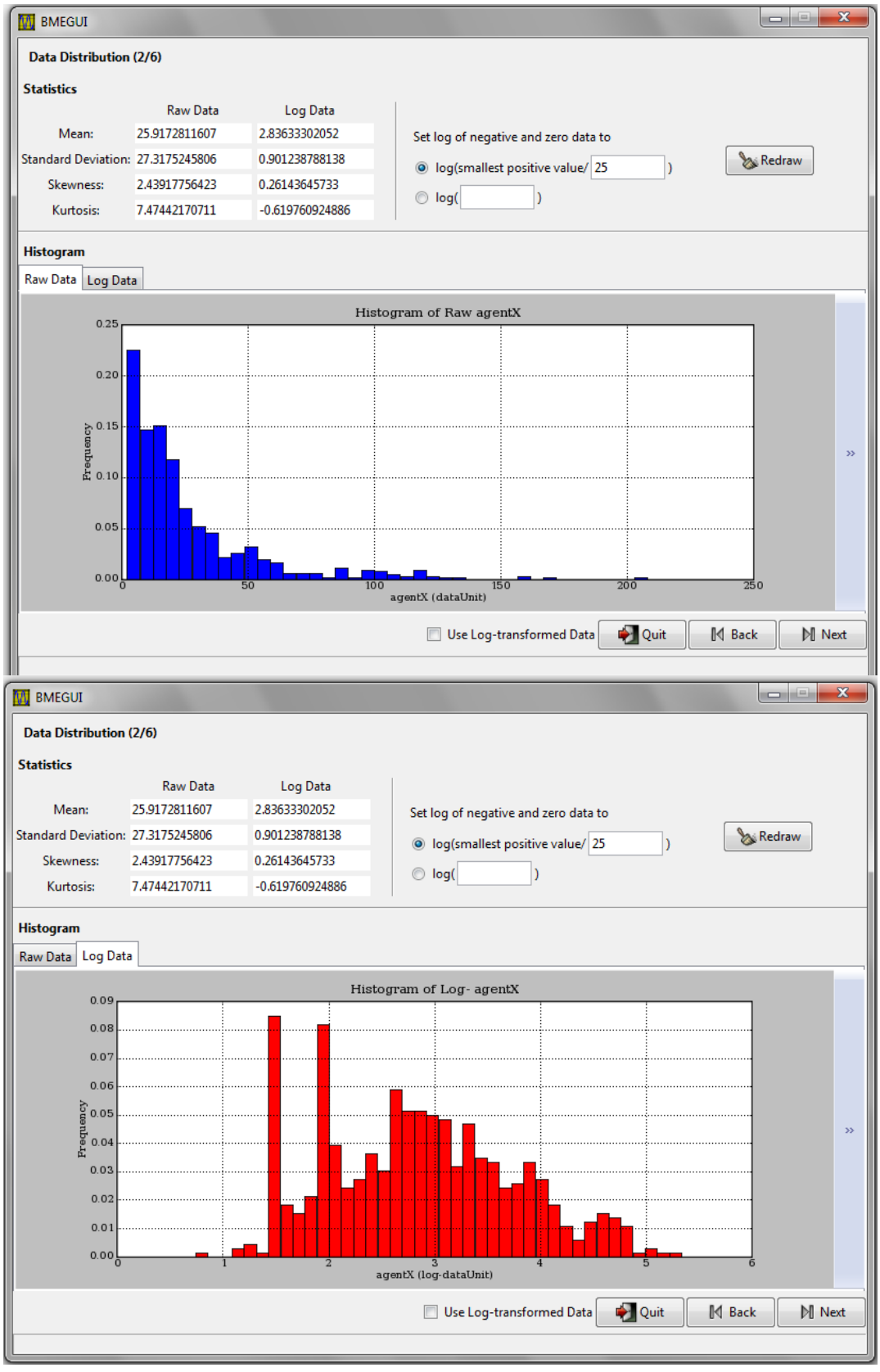


Figure 6: Dialog Box 2 (Data Distribution)

4.4.2 Data Transformation Method

Based on the basic statistics and the histogram, the user can select the data-transformation method used in the analysis. In order to use log-transformed data in the analysis, the user must check the "Use Log-transformed Data" check box, otherwise the raw data (i.e. not log-transformed data) is used (**Error! Reference source not found.**).

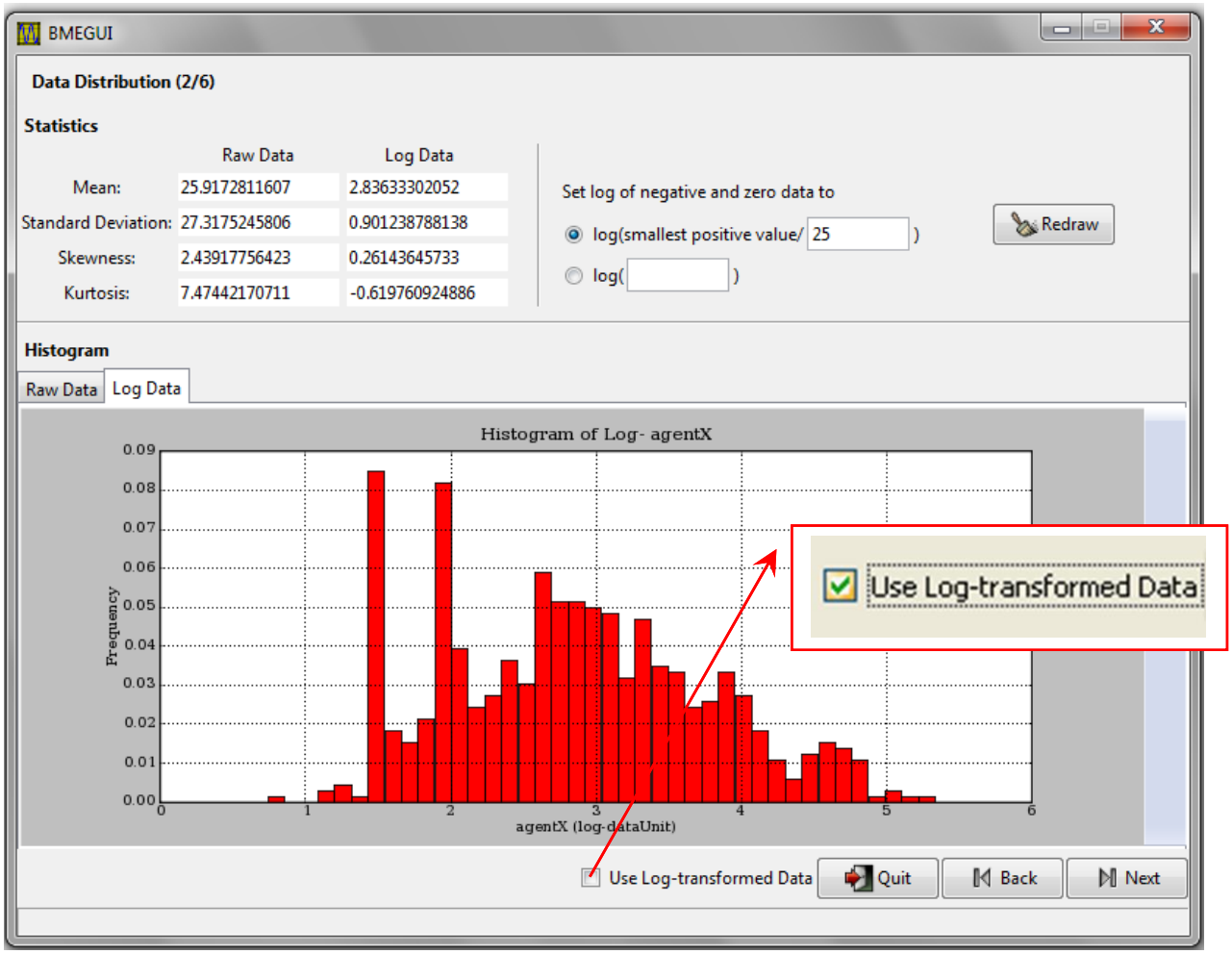


Figure 7: Use Log-transformed data

If the user selects "Use Log-transformed Data", then the histogram automatically switches to the "Log Data" tab. Similarly, if the user unselects this check box, then the histogram automatically switches to the "Raw Data" tab.

4.4.3 Log of Zero and Negative Value Setting

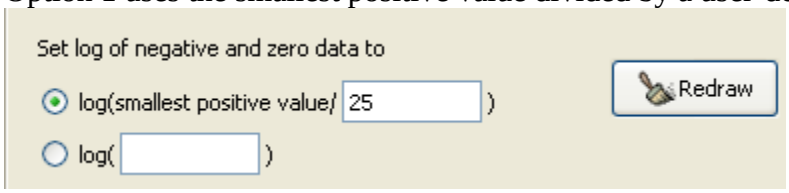
BMEGUI provides two options for dealing with the log of zero and negative values. It assigns for each zero or negative values a log-value that is either:

- ❖ The smallest strictly positive value divided by a user-defined integer, or
- ❖ The log of a user-defined value

The default setting is to use the smallest strictly positive value divided by 25. To change this setting, follow the steps described below:

1) Select the method you want to use by clicking on the corresponding radio button
2) Input the integer for the first option (**Error! Reference source not found.** a.) , or
   Input the number for the second option (**Error! Reference source not found.** b.)
3) Click on the "Redraw" button, then the basic statistics and the histogram will be updated

a. Option 1 uses the smallest positive value divided by a user-defined integer



b. Option 2uses the log of a user-defined number

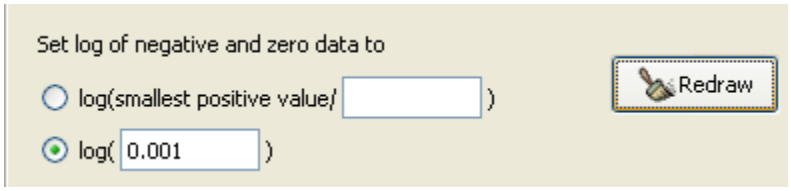


Figure 8: Settings for the log of negative and zero data values

4.4.4 Soft Data in Histogram

Since the soft data are defined in terms of their probability density function (PDF) (i.e. either the uniform or Gaussian PDF), the data have to be "hardened" before calculating their basic statistics and plotting the histogram. BMEGUI converts the soft data into hard data using the following method.

- ❖ Soft uniform data: Mid-point of lower and upper bound
- ❖ Soft Gaussian data: Mean value

"Hardened" values are also used in the following steps.

- ❖ Explanatory data analysis
- ❖ Mean trend estimation
- ❖ Experimental covariance calculation

4.5 Dialog Box 3 (Exploratory Data Analysis)

4.5.1 Basic Operation

Dialog Box 3 (Exploratory Data Analysis) shown in **Error! Reference source not ound.** is used to conduct the exploratory data analysis. This dialog box has two tabs, labeled “Temporal Evolution” and “Spatial Distribution”, respectively. BMEGUI displays the time series plot of the measurement values at each monitoring location on the “Temporal Evolution” tab, and the spatial distribution plot of the measurement values at specific times on the “Spatial Distribution” tab.

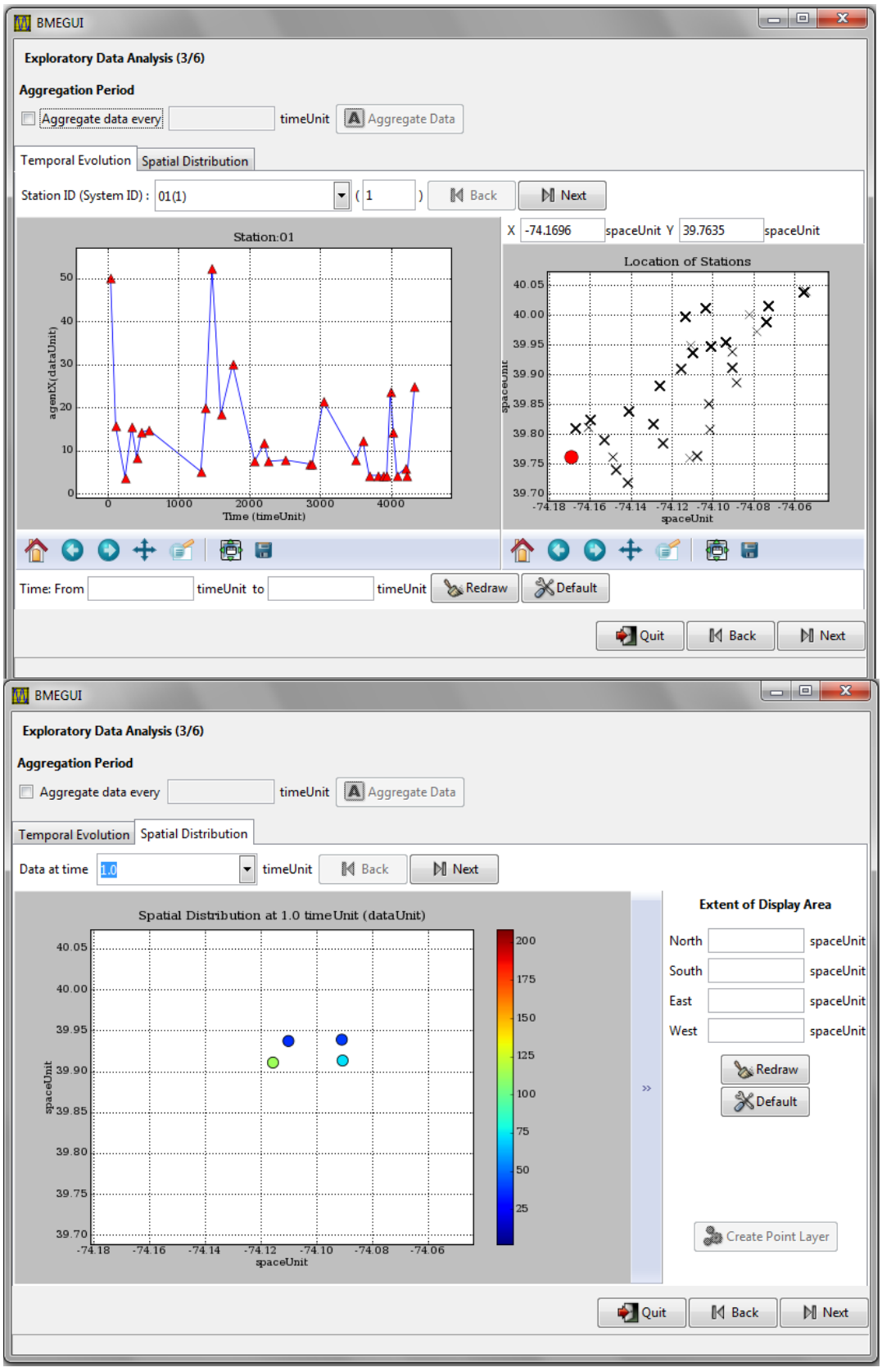


Figure 9: Dialog Box 3 (Exploratory Data Analysis)

In the “Aggregation Period” section, the user can temporarily aggregate the data using the user-defined time periods.

On the “Temporal Evolution” tab, the user can select different monitoring location of interest based on their user-defined station ID or system ID. There are three methods to select the monitoring location (Figure 9).

- Select the user-defined station ID from the dropdown menu
- Input the system ID in the entry box
- Click on the “Next” or “Back” buttons

When a new location is selected, the plot of the time series of the data available for that location is automatically updated.

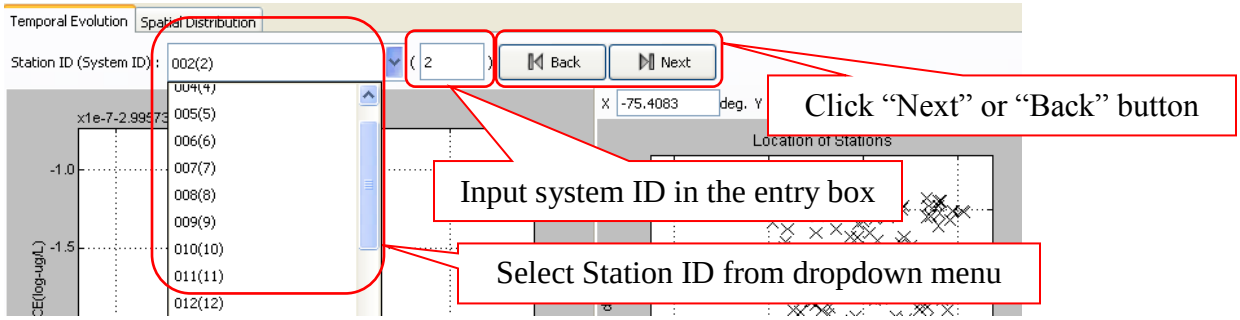


Figure 9: “Temporal Evolution” tab - Three methods to select the monitoring location

Similarly, on the “spatial distribution” tab, the user can select specific times (for which to create spatial plots of the available data) using the dropdown menu or the “Next” or “Back” buttons (**Error! Reference source not found.**).

Figure 10: “Spatial Distribution” tab - Methods to select specific times

4.5.2 Data Aggregation

In Dialog Box 3, the user can aggregate the data temporally using user-defined aggregation time periods. When the data is aggregated, all the measurement values within

a given aggregation period are treated as if they are sampled at the same time (**Error! eference source not found.**).

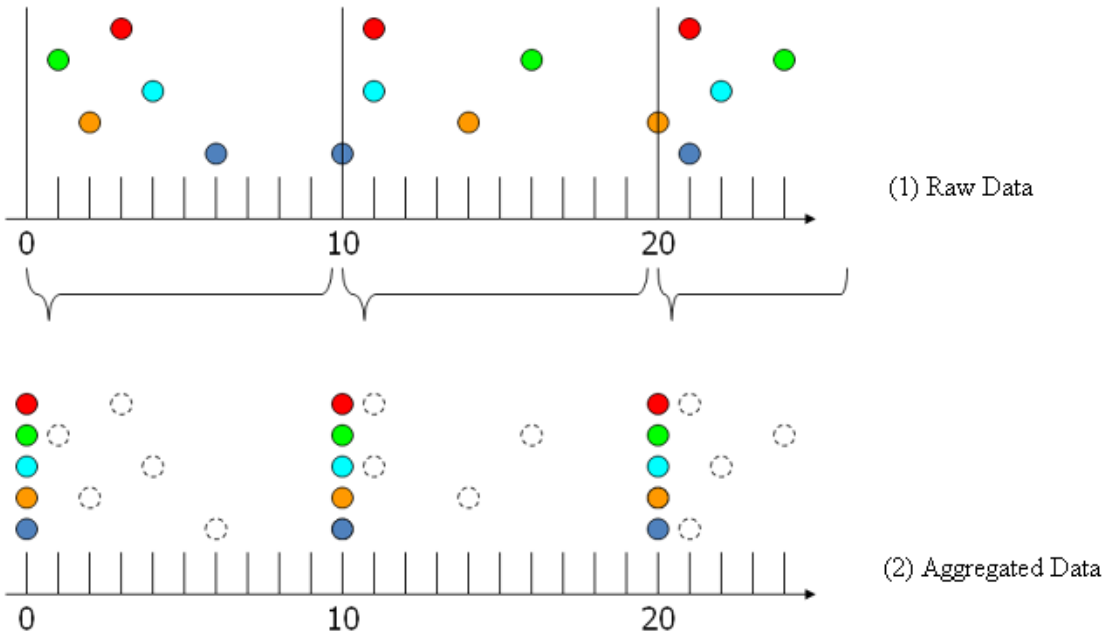


Figure 11: Example of data aggregation with 10 time-unit aggregation period. (1) raw data and (2) aggregated data

The aggregated data is used to create the spatial distribution plots in Dialog Box 3, for mean trend analysis in Dialog Box 4, and to obtain the experimental covariance in Dialog Box 5.

To aggregate the data, follow the steps described below.

1) Check the box (Aggregate data every …), then the entry box "Aggregate Data" button will be activated (**Error! Reference source not found.** (1) and (2)).
2) Enter the aggregation period (**Error! Reference source not found.** (3)) in the ntry box.
3) Click the "Aggregate Data" button, then the data will be aggregated and the button will be deactivated (**Error! Reference source not found.** (4)).
4) To go back to the non-aggregated data, uncheck the box (Aggregate data every…).

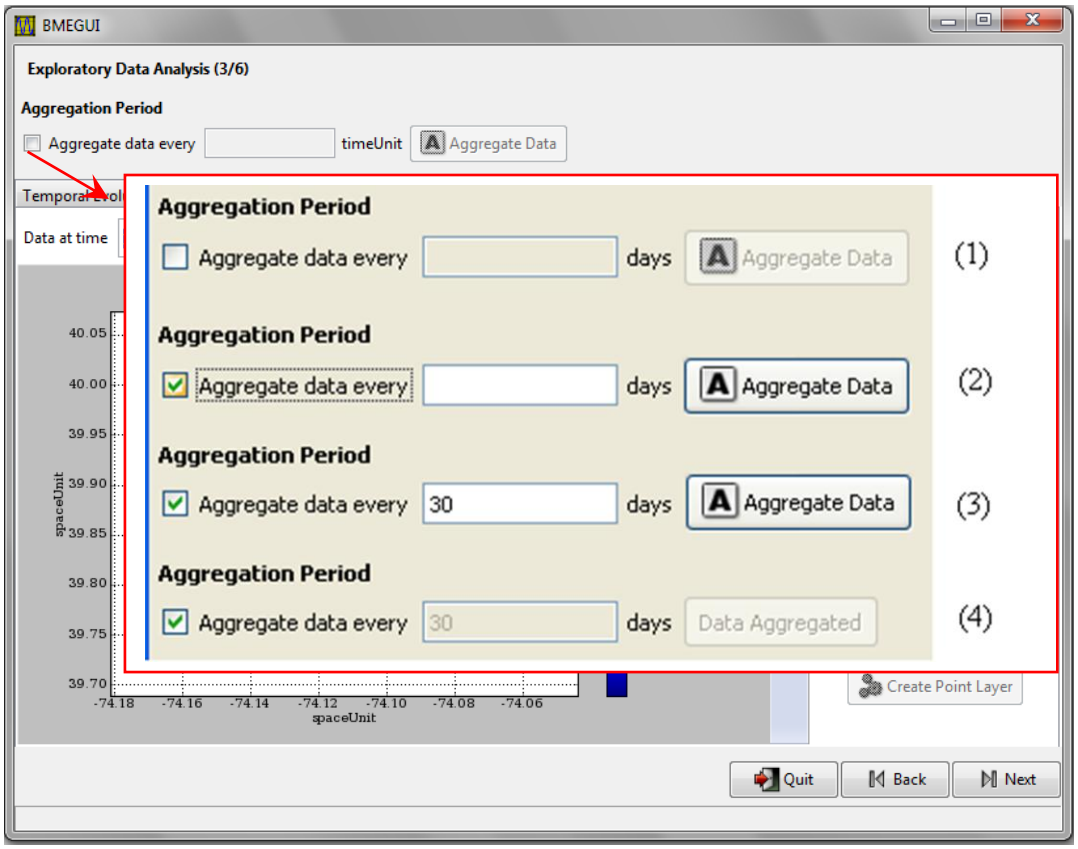


Figure 12: Data aggregation

4.5.3 Create Point Layer File

The user can create ArcGIS, MATLAB (and other software) compactible output file. BMEGUI pop ups a window asking file and directory names where data files have to be saved. Such file is saved by BMEGUI as CSV (*.csv) file regardless of extension supplied by users.

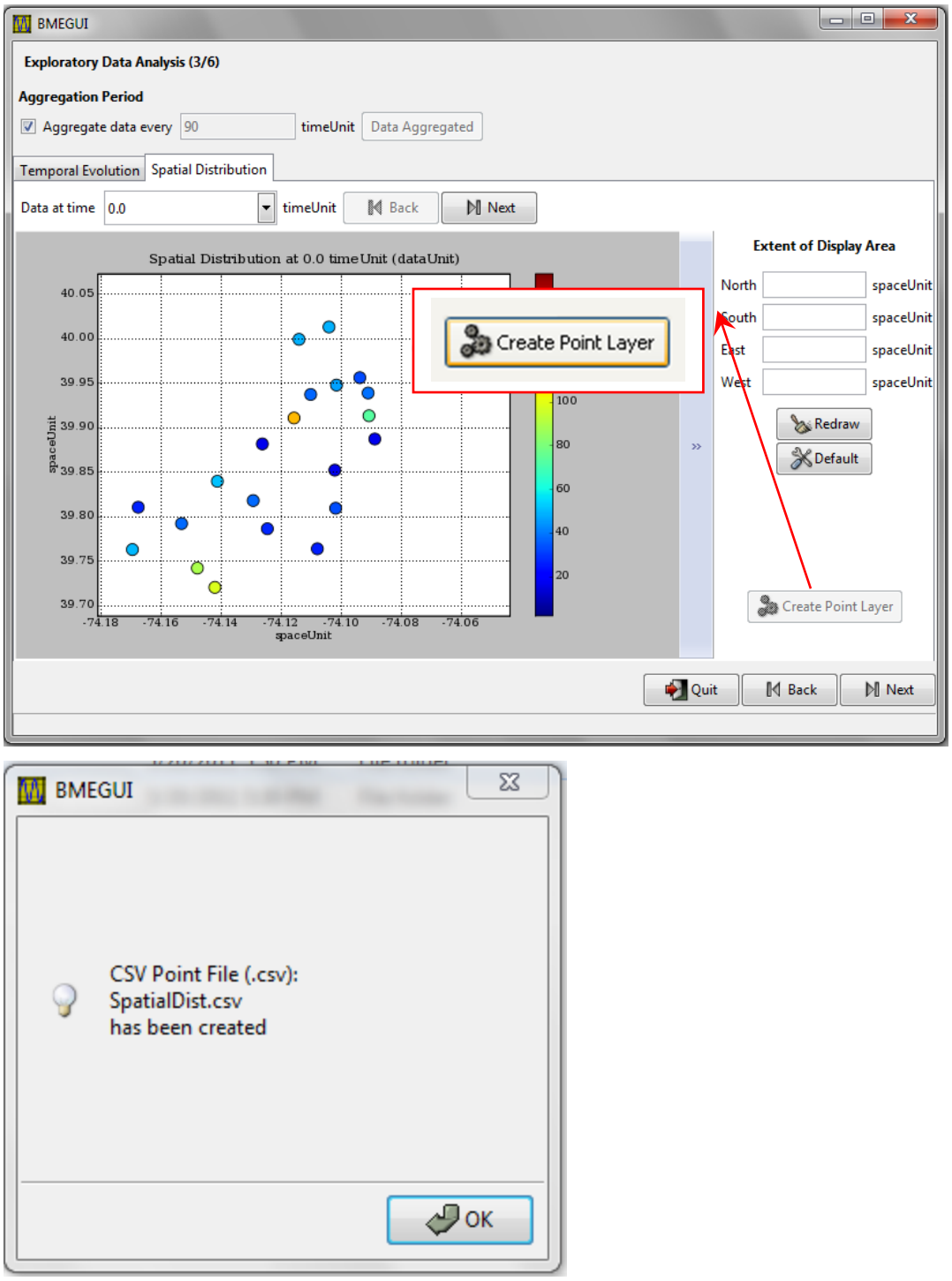


Figure 13: The "Create Point Layer" button and the message box

4.6 Dialog Box 4 (Mean Trend Analysis)

4.6.1 Basic Operation

Dialog Box 4 (Mean Trend Analysis) shown in **Error! Reference source not found.** llows the users to explore whether the data exhibits a global trend across space and time.

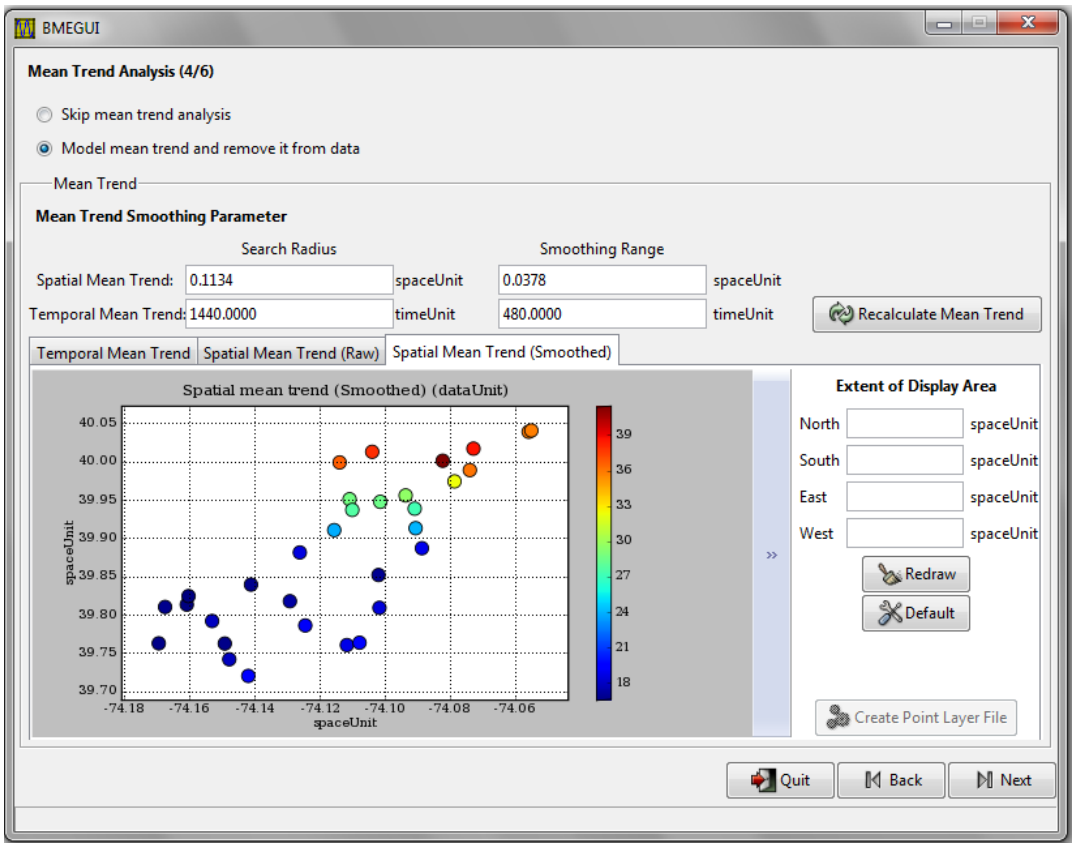


Figure 14: Dialog Box 4 (Mean Trend Analysis)

A global mean trend is a function of space and time that describes consistent patterns in the data, i.e. it describes where or when the data seems to be consistently higher or consistently lower than the mean. The word "global" emphasizes that this trend applies globally to the whole space/time domain encompassing all the available data. Dialog Box 4 displays the global mean trend, and the user must decide whether this global trend should be used in further analysis. If the global mean trend is used, then it is removed from the data, yielding residual values (i.e. data minus global trend) that are then used in the ensuing analysis (i.e. for the covariance analysis and BME estimation). Hence, the goal of the global mean trend should be to produce residuals that are as homogeneous (i.e. without spatial trend) and stationary (i.e. without temporal trend or drift) as possible. As a default setting, BMEGUI does not calculate the mean trend nor does it remove it from the data.

BMEGUI assumes that the global mean trend $m_{st}(\boldsymbol{s},t)$, where $\boldsymbol{s}$ denotes the spatial coordinate and t is time, is a space/time additive separable function, i.e. that it has the following form

$$m(\boldsymbol{s},t) = m_{ss}(\boldsymbol{s}) + m_{ts}(t)$$

where $m_{ss}(\boldsymbol{s})$ is the spatial component smoothed over space and $m_{ts}(t)$ is the temporal component smoothed over time (also called the temporal drift). BMEGUI first averages the measurements at each monitoring sites to obtain values for the raw spatial mean m_s, and then it applies an exponential spatial filter to these raw spatial mean values to obtain a spatial component m_{ss} that is smoothed over space. Conversely, BMEGUI first

averages the measurements for each monitoring time event (or each aggregated time periods if the data has been time aggregated) to obtain values for the raw temporal mean m_t, and then it applies an exponential temporal filter to these raw average values (minus their overall average) to obtain a temporal component m_{ts} that is smoothed over time.

This dialog box has three tabs, namely the "Temporal Mean Trend" tab showing both the raw temporal average values m_t and the temporal trend component m_{ts} smoothed over time, the "Spatial Mean Trend (Raw)" tab showing the raw spatial average values m_s, and the "Spatial Mean Trend (Smoothed)" tab showing the spatial mean trend component m_{ss} smoothed over space.

To calculate the mean trend using the method described above and remove it from the data, click on the "Model mean trend and remove it from the data" radio button. Then BMEGUI calculates the mean trend using the default parameter (**Error! Reference ource not found.**).

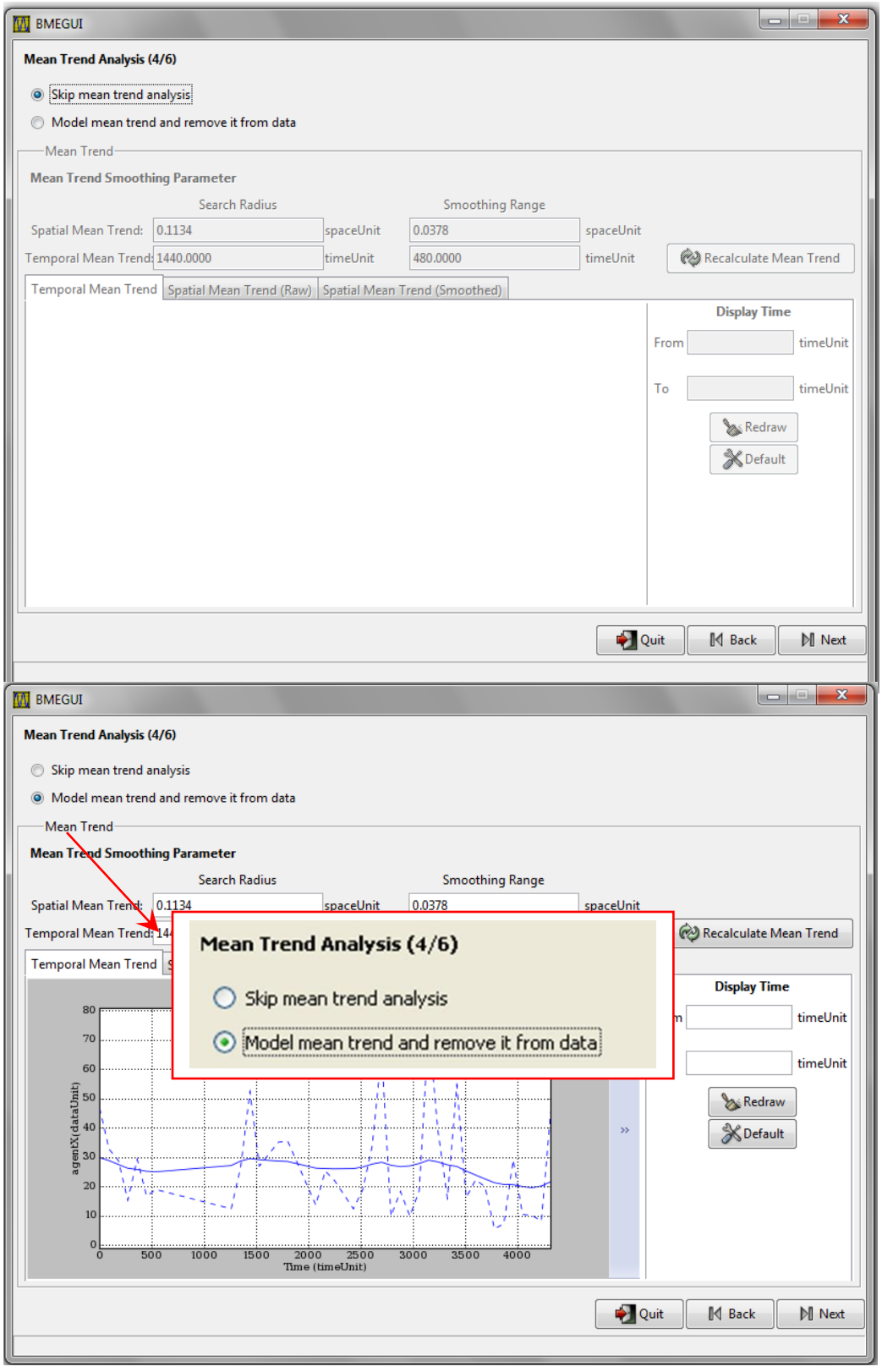


Figure 15: Calculating the global mean trend and removing it from the data

4.6.2 Calculate Mean Trend Using User-defined Parameters

The user can calculate the global mean trend using user-defined parameters. There are two parameters which are used to control the spatial exponential filter used to smooth the raw spatial averages m_s in order to obtain the smoothed spatial trend m_{ss}:

- The "Spatial Search Radius", corresponding to the radius of the spatial neighborhood used select points for the spatial exponential filter
- The "Spatial Smoothing Range", corresponding to the range of the spatial exponential function.

Similarly, there are two parameters which are used to control the temporal exponential filter used to smooth the raw temporal averages m_t in order to obtain the smoothed temporal trend m_{ts}:

- The "Temporal Search Radius", corresponding to the radius of the temporal neighborhood used to select points for the temporal exponential filter
- The "Temporal Smoothing Range", corresponding to the range of the temporal exponential function.

To calculate the mean trend, input these four parameters in the "Mean Trend Smoothing Parameter" section. Then, click on the "Recalculate Mean Trend" button. The plots of smoothed temporal and spatial mean trends will be updated (**Error! Reference source ot found.**). In order to make the spatial or temporal trend smoother, increase the corresponding two parameter values, and recalculate the trend. Conversely to obtain a trend that is less smooth (i.e. that follows more closely the raw averages), decrease the parameters values and recalculate the trend.

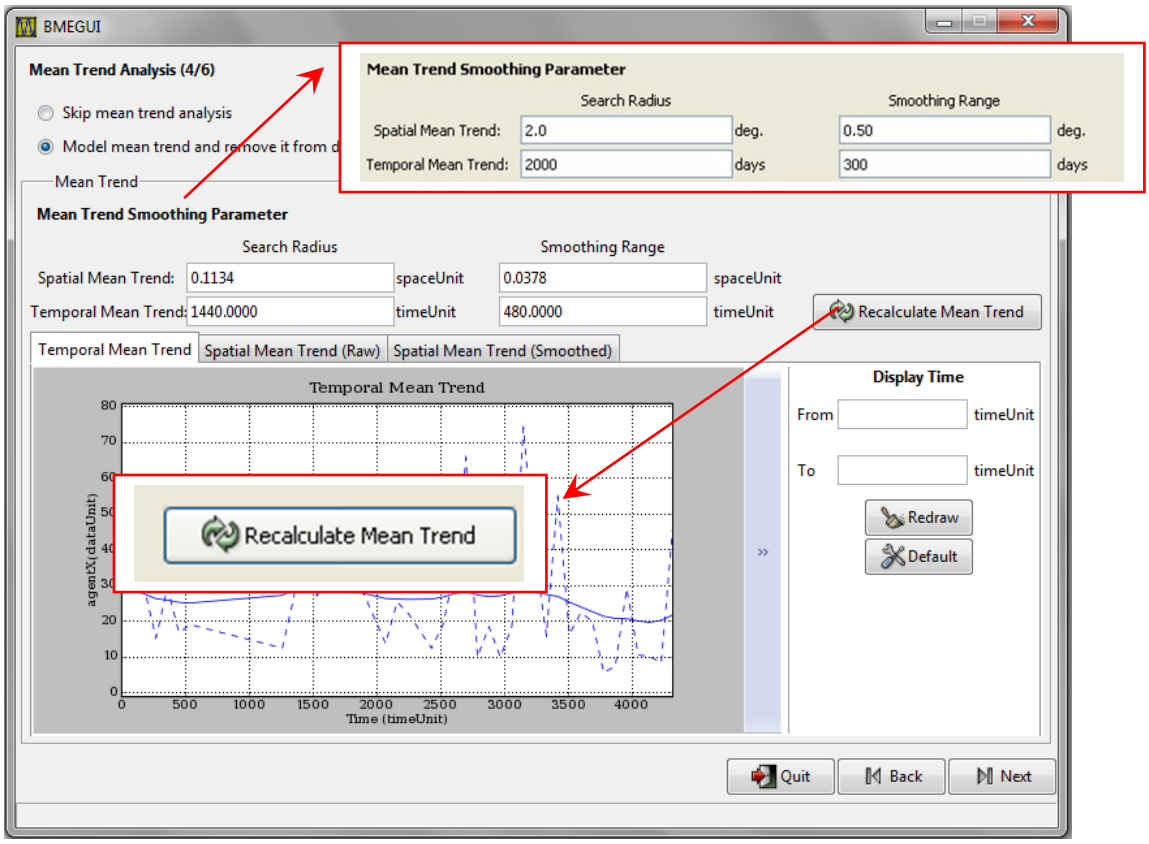


Figure 16: The mean trend smoothing parameters and the "Recalculate Mean Trend" button

4.6.3 Create Point Layer File

Similarly to with the spatial distribution plot in Dialog Box 3 (see section 4.5.3), the user can create a point layer file of the raw and smoothed spatial mean trend. To create this point layer file, click on the "Create Point Layer" button. Then a message box will appear asking for file name and directory name where file has to save. After user provide file name and directory and click on OK button, another message box will appear saying the name of the point layer file has been created.

4.7 Dialog Box 5 (Space/Time Covariance Analysis)

4.7.1 Basic Operation

Dialog box 5 (Space/Time Covariance Analysis) shown in **Error! Reference source not ound.** is used to calculate the spatial and temporal components of the covariance of the data (or of its residual if the mean trend was removed from the data). The data (or its residual) are assumed to be homogeneous and stationary, which implies that the covariance between two space/time points $\boldsymbol{p}=(\boldsymbol{s},t)$ and $\boldsymbol{p}'=(\boldsymbol{s}',t')$ is only a function of the spatial lag (i.e. the spatial distance) $r=\|\boldsymbol{s}-\boldsymbol{s}'\|$ and time lag (i.e. the time difference) $\tau=|t-t'|$ between these two space/time points. Hence the covariance $c(\boldsymbol{p},\boldsymbol{p}')$ between points $\boldsymbol{p}$ and $\boldsymbol{p}'$ can be written as

$$c(\boldsymbol{p},\boldsymbol{p}') = c(r=\|\boldsymbol{s}-\boldsymbol{s}'\|,\ \tau=|t-t'|),$$

where r is the spatial lag and τ is the temporal lag.

There are two steps in modeling the covariance. First we need to estimate the covariance value for different spatial and temporal lags. We call these estimated values the "experimental covariance" values. Then we need to fit a permissible covariance model to the experimental covariance values.

In order to simplify the visual representation of the fitting of the covariance model $c(r,\tau)$ to the experimental covariance values, Dialog Box 5 shows the 2-dimensional covariance function in terms of two distinct one-dimensional plots. The first plot is shown on the "Spatial Component" tab (**Error! Reference source not found.**a), and it is a plot of the ovariance $c(r,\tau=0)$ with respect to the spatial lag r for τ=0. The second plot is shown on the "Temporal Component" tab (**Error! Reference source not found.**b), and it is a plot f the covariance $c(r=0,\tau)$ with respect to the temporal lag τ for r=0.

On the "Spatial Component" tab, the experimental values of $c(r,\tau=0)$ are estimated for a set of user-defined spatial lags r plus/minus a corresponding set of spatial lag tolerances dr. For example if the spatial lags are r={3, 6} and the corresponding spatial tolerances are dr={1, 2}, then the experimental covariances on the Spatial Component tab will be estimated for τ=0 and r=3+/-1 (i.e. using all pairs of points with a temporal lag of zero and a spatial lags between 2 and 4), and for τ=0 and r=6+/-2 (i.e. using all pairs of points with a temporal lag of zero and a spatial lags between 4 and 8). Conversely on the "Temporal Component" tab, the experimental values of $c(r=0,\tau)$ are estimated for a set of user-defined temporal lags τ plus/minus a corresponding set of spatial lag tolerances $d\tau$.

The next section explains how to modify the spatial and temporal lags in Dialog Box 5 to calculate the experimental covariance values, and the following section explains how to use Dialog Box 5 to fit the covariance model on to the experimental covariance values.

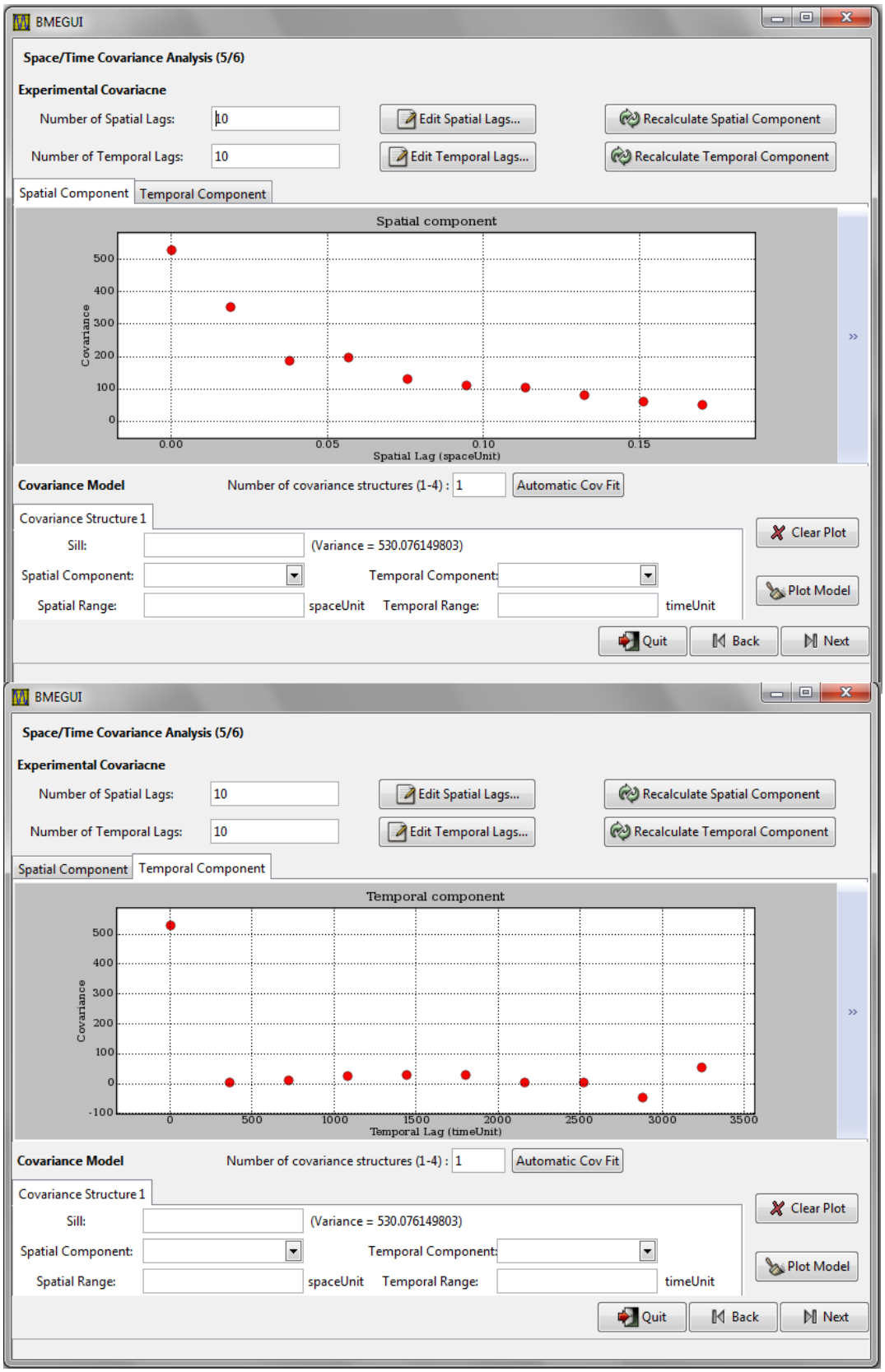


Figure 17: Dialog Box 5 (Space/Time Covariance Analysis)

4.7.2 Calculate Experimental Covariance

There are two methods to set the spatial and temporal lags used to calculate the experimental covariance values. One is to simply set the number of lags used, in which case BMELIB uses equidistant lags and a constant (identical) lag tolerance for each lag. By default, BMEGUI automatically calculates the experimental covariance using 10 equidistance lags. The other method is to enter each lag and corresponding lag tolerance individually, which offers the flexibility that the lags need not be equidistant.

To modify the number of the lags, follow these steps (**Error! Reference source not ound.**):

1) Input the number of the spatial or temporal lags you would like to use in the entry box. In this case BMELIB sets equidistant spatial lags from 0 to half of the maximum distance between data points, and equidistant temporal lags from 0 to half of the maximum time difference between data points
2) Click on the "Recalculate Spatial Component" or "Recalculate Temporal Component" buttons.
3) The experimental covariance plot will be updated.

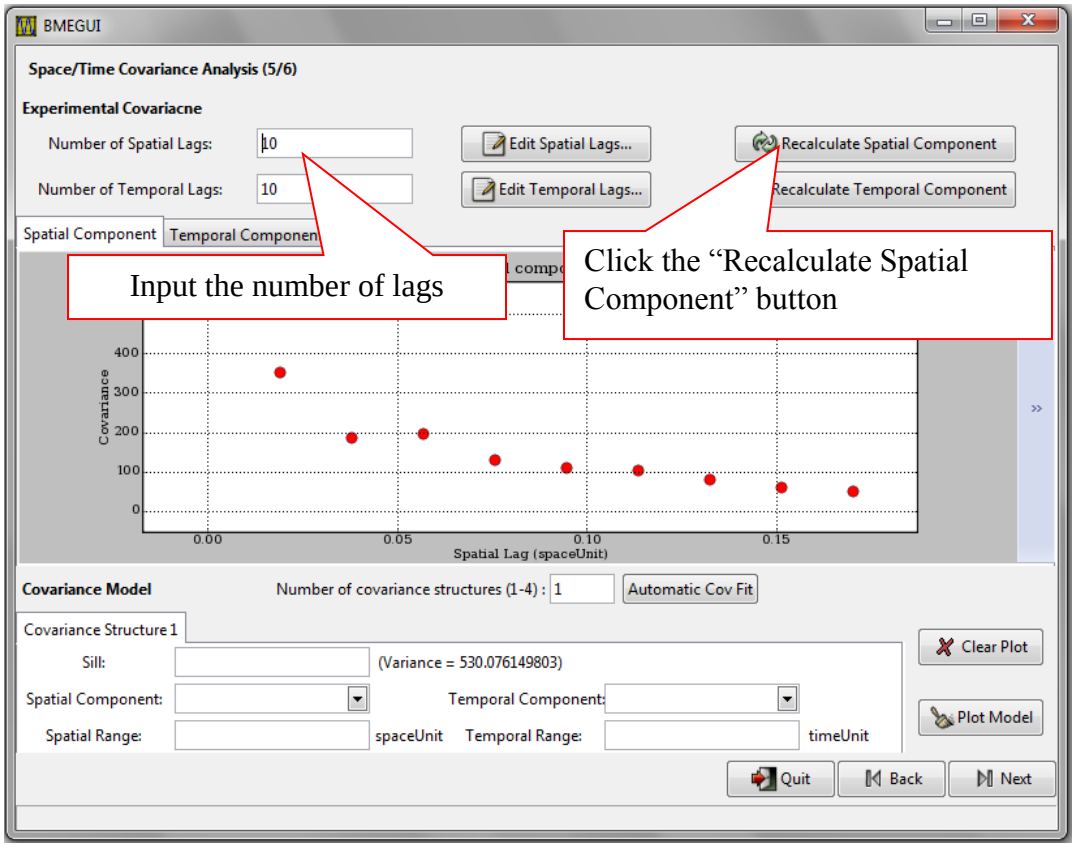


Figure 18: Calculating experimental covariance by modifying the number of the lags

To directly enter the lags and their corresponding lag tolerances, follow these steps (**Error! Reference source not found.**):

1) Click on the "Edit Spatial Lags…" or "Edit Temporal Lags…" buttons, and then a dialog box will appear.
2) Input the lags values (e.g. 0.0,0.019,0.038,0.057,0.076,0.094,0.113,0.132, 0.151,0.170) and a corresponding number of lag tolerances (e.g. 0.0,0.009,0.009, 0.009,0.009,0.009,0.009,0.009,0.009,0.009) in the entry box. Use commas (,) to delimit values.
3) Click on "OK"
4) The experimental covariance plot will be updated.

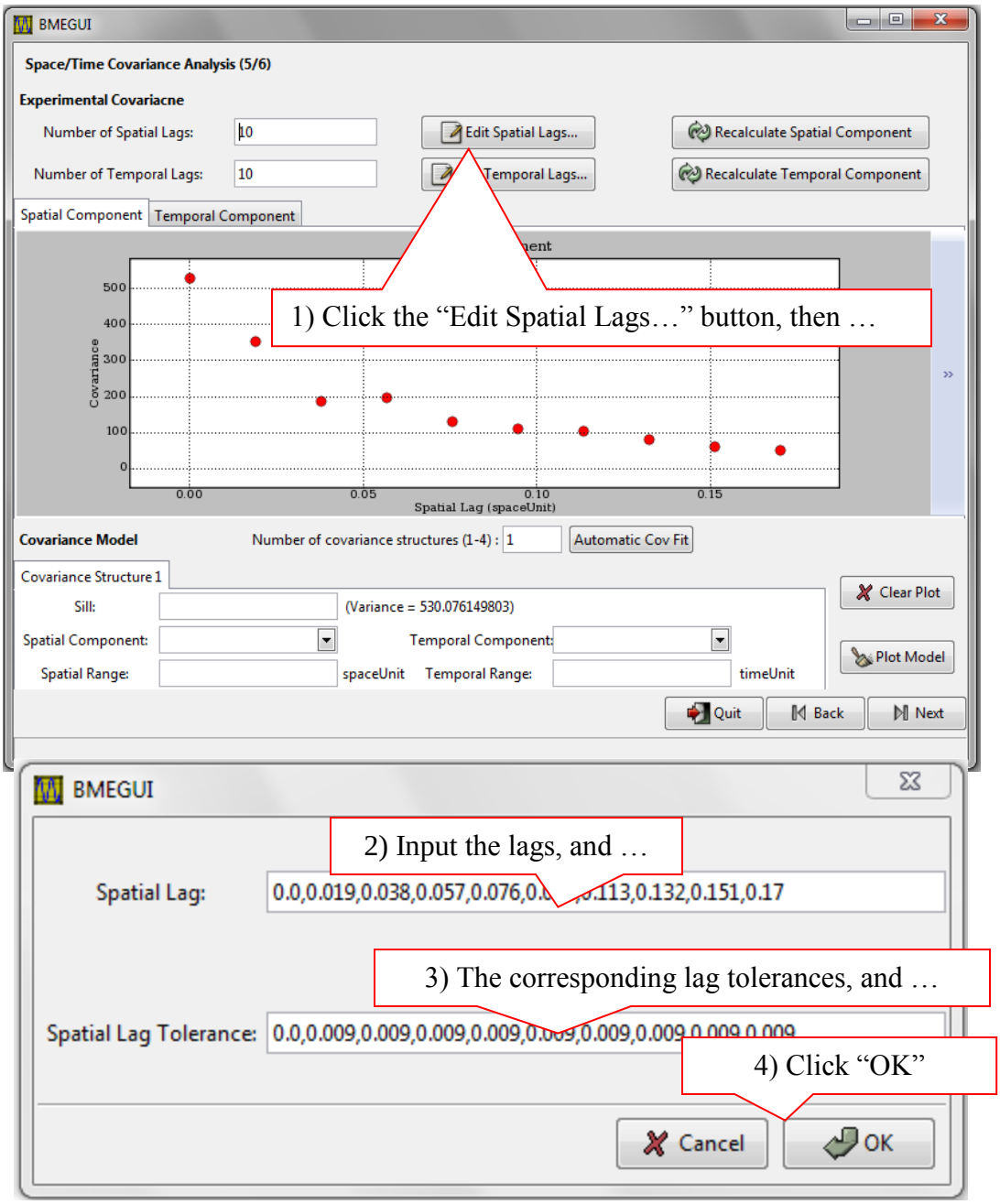


Figure 19: Calculating experimental covariance values by directly entering the lags and the lag tolerances

4.7.3 Covariance Model

The user must select a space/time covariance model that fits the experimental covariance values. BMEGUI lets the user select that model among the large class of space/time covariance models given by the following equation

$$c(r,\tau) = \sum_{i=1}^{N} c_{0i}\, c_{ri}(r)\, c_{ti}(\tau)$$

where N is the number of covariance structures, c_{0i} is the variance contribution (or "*sill*") of the *i-th* covariance structure, and $c_{ri}(r)$ and $c_{ti}(\tau)$ are permissible functions representing the spatial and temporal components, respectively, of the *i-th* covariance structure. BMEGUI supports up to four structures, i.e. $N \leq 4$. The permissible covariance functions for the spatial components $c_{ri}(r)$ include the following

- Exponential: $c_{ri}(r) = \exp(-\frac{3r}{a_{ri}})$
- Gaussian: $c_{ri}(r) = \exp(-\frac{3r^2}{a_{ri}^2})$
- Spheroidal: $c_{ri}(r) = \left(1-\frac{3}{2}\frac{r}{a_{ri}}-\frac{1}{2}\frac{r^3}{a_{ri}^3}\right)$
- Holecos: $c_{ri}(r) = \cos(\pi r / a_{ri})$
- Holesin: $c_{ri}(r) = \left(\frac{\sin(\pi r / a_{ri})}{\pi r / a_{ri}}\right)$

and similar ones are available for the temporal component $c_{ti}(\tau)$. Generally a_{ri} and a_{ti} are called the "*spatial range*" and the '*temporal range*", respectively, of the *i-th* structure of the covariance function.

It can be noted that each of the functions used for the spatial and temporal components take a value of 1 for a lag of zero, i.e. $c_{ri}(0)=1$ and $c_{ti}(0)=1$, $i=1,\ldots,N$. Since by definition the variance of the covariance model (also called the "*model variance*") is obtained by calculating the model covariance at a spatial and time lags of zero, it follows that the model variance is given by $\sum_{i=1}^{N} c_{0i}$ because $c_{ri}(0)=1$ and $c_{ti}(0)=1$, $i=1,\ldots,N$. Since the model variance should represent the variance of the data, then the user should select the sills c_{0i}, $i=1,\ldots,N$, such that their sum is approximately equal to the variance of the data. In order to help with this constraint, BMEGUI displays the variance of the data in Dialog Box 5 (see 'Variance= xxxx "in **Error! Reference source not found.**).

To select and plot a covariance model, follow these steps (**Error! Reference source not ound.**).

1) Input the number N of the covariance structures desired (making sure that $1 \leq N \leq 4$).
2) Input the sill c_{oi} of the *i-th* covariance structure. Keep in mind that the sum of the sills should be equal to the variance of the data, which is displayed on the right side of the entry box.
3) Select the functions used to model the spatial and temporal components of the *i-th* covariance structure using the dropdown menus.
4) Input the value for the spatial range and temporal range of the *i-th* covariance structure.

5) Repeat steps from 2) to 4) for each covariance structure.
6) Click on the “Plot Model” button to plot the covariance model.
7) The covariance model (shown as a plain line) should fit the experimental covariance values (show as markers). Repeat steps 1) to 6) until the covariance model fits well with the experimental covariance values.
8) Click on the “Clear Plot” at any time during step 7) to clear the different models previously plotted.

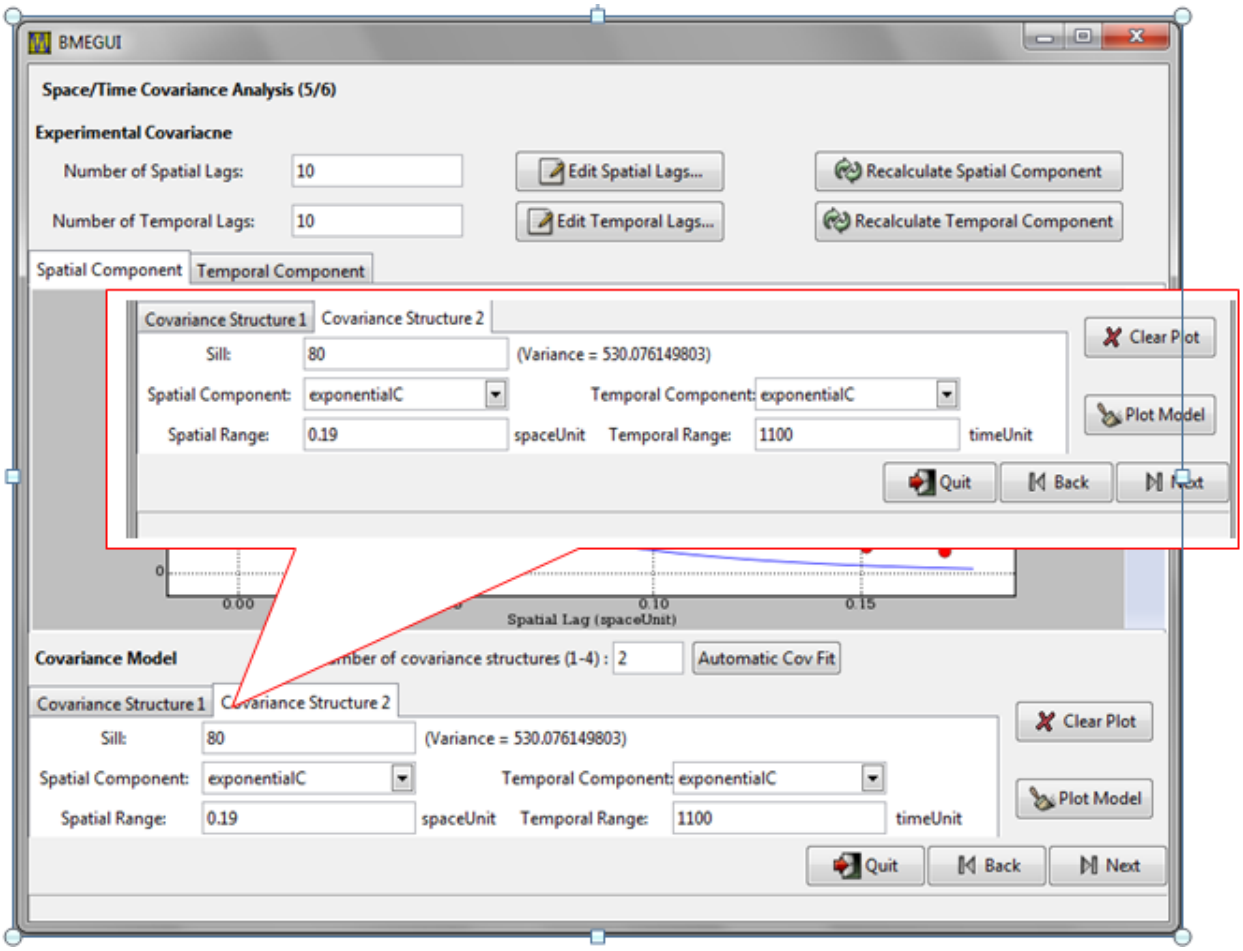


Figure 20: Covariance model parameter settings

4.7.4 Automatic Covariance Model fitting

You can also fit covariance model by clicking on ‘Automatic Cov Fit’ button as in figure 22 below. However, automatic fitting may not be a good fit for a specific problem since BMEGUI fits a single covariance model rather than a composite model. In many situations, composite covariance model of space-time separable models may be best suitable and fit model which can be done as described in previous section.

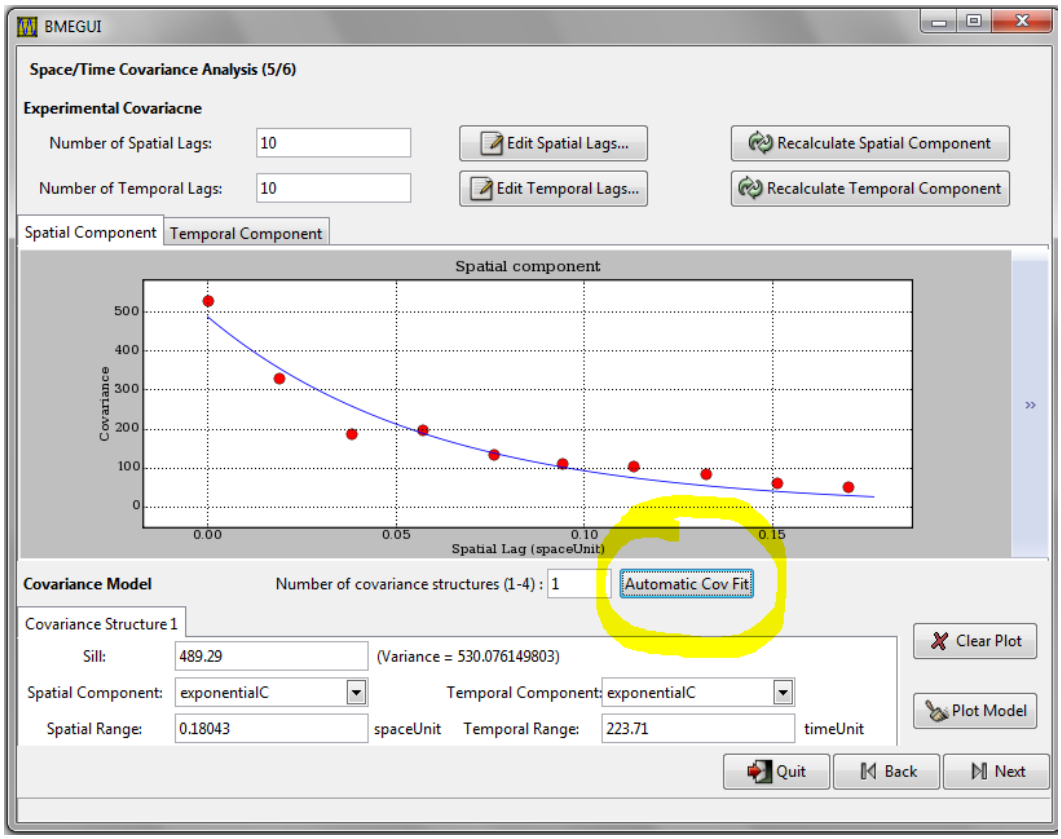
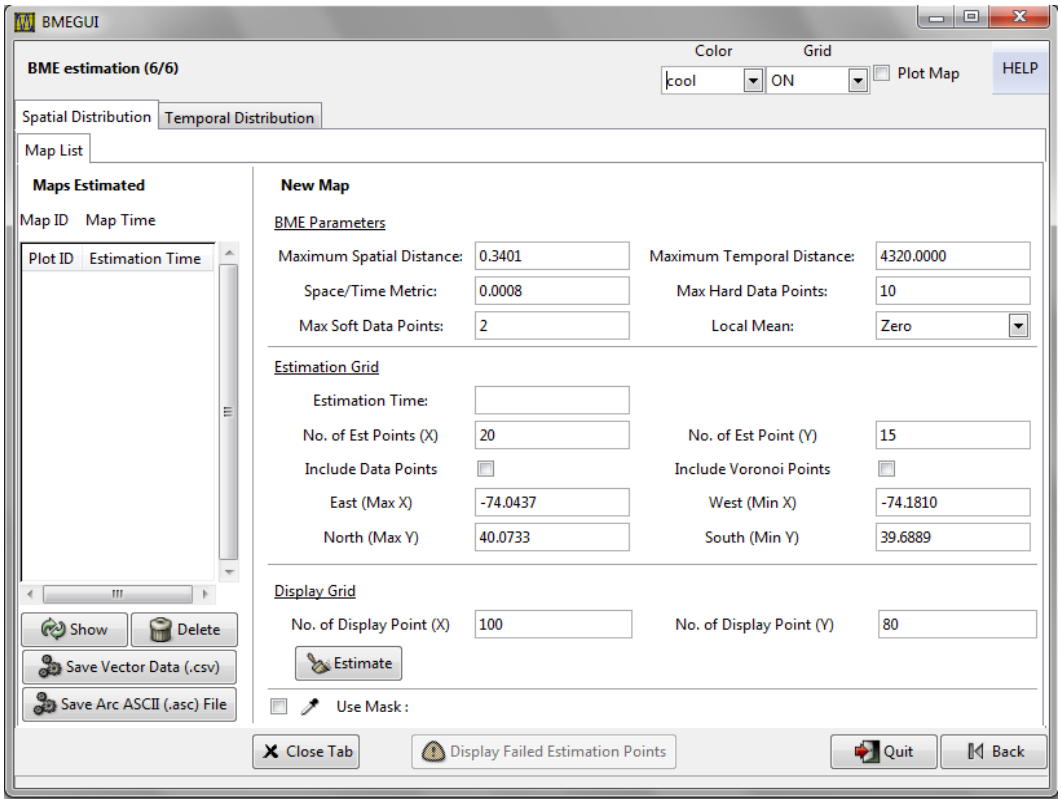


Figure 21: Automatic covariance model fitting

4.8 Dialog Box 6

4.8.1 Basic Operation

Dialog Box 6 (BME Estimation) shown in **Error! Reference source not found.** is used o calculate BME estimated values. Dialog Box 6 has two tabs, the "Spatial Distribution" tab and the "Temporal Distribution" tab. The "Spatial Distribution" tab is used to create maps of the BME mean estimates and the BME error variance at specific times of interest. The "Temporal Distribution" tab is used to create plots (also called "time series") of the BME mean estimate and BME error variance as a function of time for specific monitoring locations of interest.



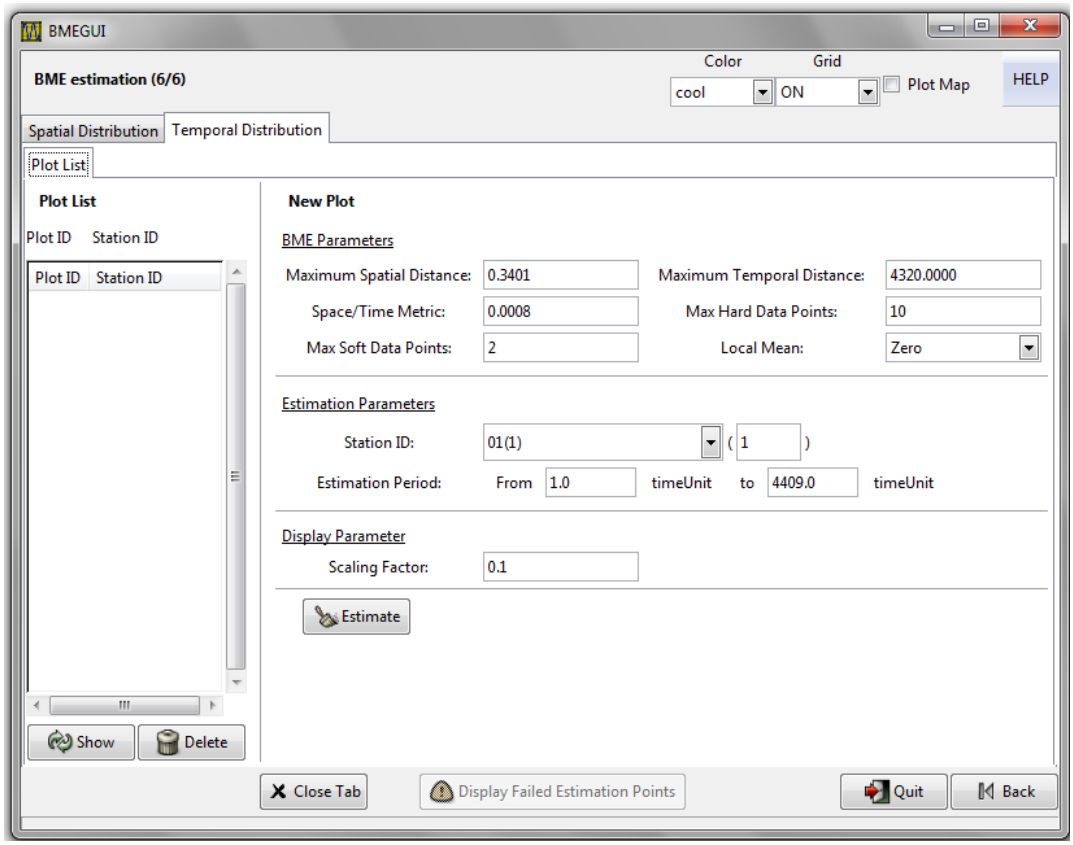


Figure 22: Dialog Box 6 (BME Estimation)

4.8.2 BME Parameters

The user needs to specify the following six BME estimation parameters to obtain BME estimated values both on "Spatial Distribution" tab and on "Temporal Distribution" tab.

- Maximum Spatial Distance: The maximum spatial distance between an estimation location and data locations.
- Maximum Temporal Distance: The maximum temporal lag between an estimation location and data locations
- Space/Time Metric: A parameter that is used to calculate the space/time distance. The space/Time distance is obtained as (Spatial distance) + (Space/Time Metric) * (Temporal distance)
- Max Hard Data Point: The maximum number of hard data values used in the estimation
- Max Soft Data Point: The maximum number of soft data values used in the estimation
- Local Mean: Order of the polynomial used to model the mean trend (or drift) along the spatial and temporal axes within the neighborhood of the estimation point. The default setting is "Zero", which will use a mean trend of zero and corresponds to simple kriging. "Constant" will use a constant local drift, which corresponds to ordinary kriging applied locally around the estimation point. "Linear" will use a local drift that varies linearly along the spatial and temporal axes. "Quadratic" will use a polynomial of order 2, etc. Generally Order>1 corresponds to universal kriging applied locally around the estimation point.

The values of these parameters are displayed in the "BME parameters" section in each tab. BMEGUI automatically displays default BME parameters, however; the user can modify these parameters (**Error! Reference source not found.**).

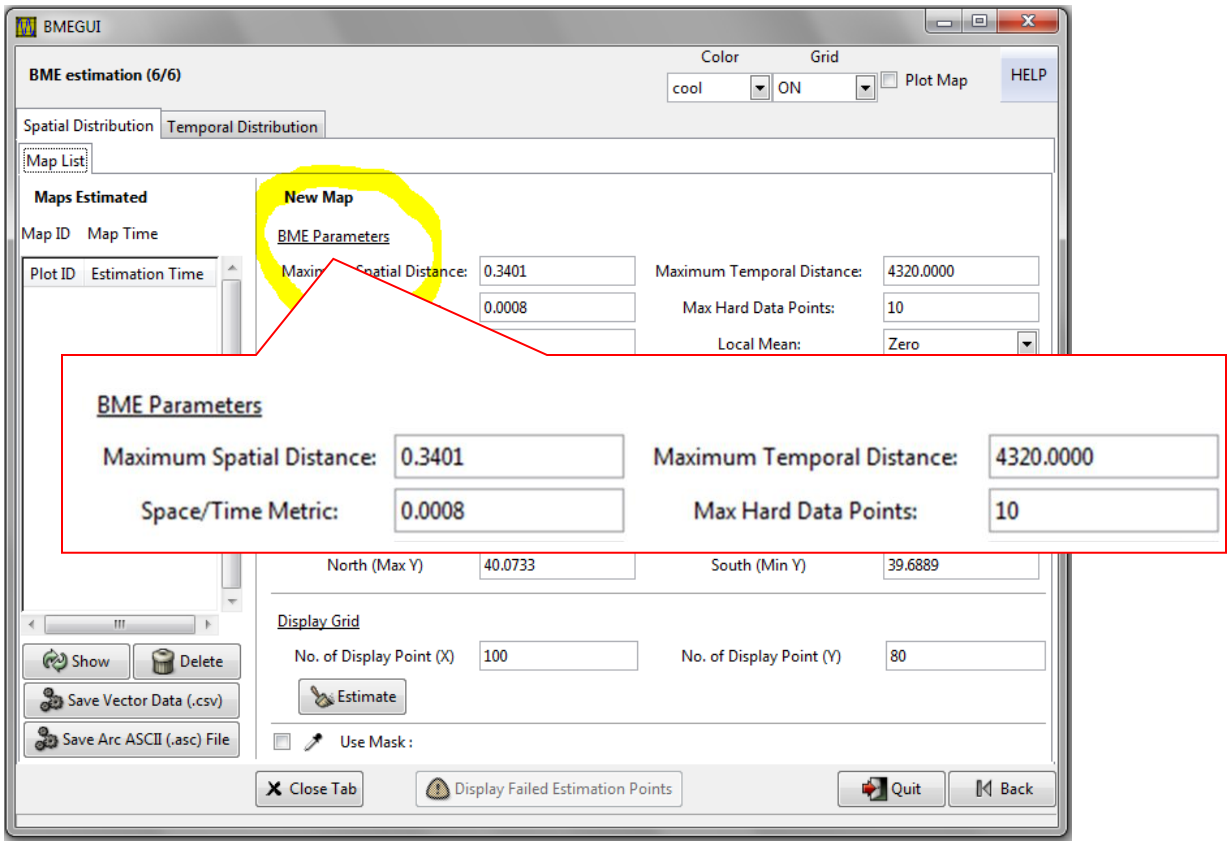


Figure 23: BME Parameters

4.8.3 Estimation Parameters (Spatial Distribution)

In order to obtain maps of BME estimates, the user needs to specify the "Estimation Grid" parameters and the "Display Grid" parameters. To obtain a map, first BMEGUI creates an "Estimation Grid" consisting of estimation nodes distributed across space within a user-defined rectangle area, and calculates the BME estimates at these estimation nodes. Then BMEGUI creates a "Display Grid" consisting of nodes distributed over a fine regular grid within the defined rectangle area, and linearly interpolates the BME estimates at estimation nodes onto the display regular grid. This two-step process speeds up the creation of the map.

In the "Estimation Grid" section, the user can specify the following parameters (**Error! eference source not found.**).

- Estimation Time: The time of interest for which to produce the BME map. There is no default value for this field.
- Number of Estimation Points (X) and (Y): The number of estimation grid points along the X-axis and Y-axis
- Area of Estimation Grid: Boundaries of the rectangle where the estimation grid is created. The user can specify the following four boundaries: East(Max X), West(Min X), North(Max Y) and South(Min Y)

In addition, the user can include to the estimation grid all the monitoring locations, as well as the set of Voronoi points constructed from these monitoring locations. Adding these points will increase the computation time, but it will lead to maps with finer spatial details. To include these points, check the "Include Data Points" box or "Include Voronoi Points" box in "Estimation Grid" section.

In the "Display Grid" section, the user can specify the number of display grid points along the X-axis and Y-axis (**Error! Reference source not found.**). A regular grid is hen constructed using these settings.

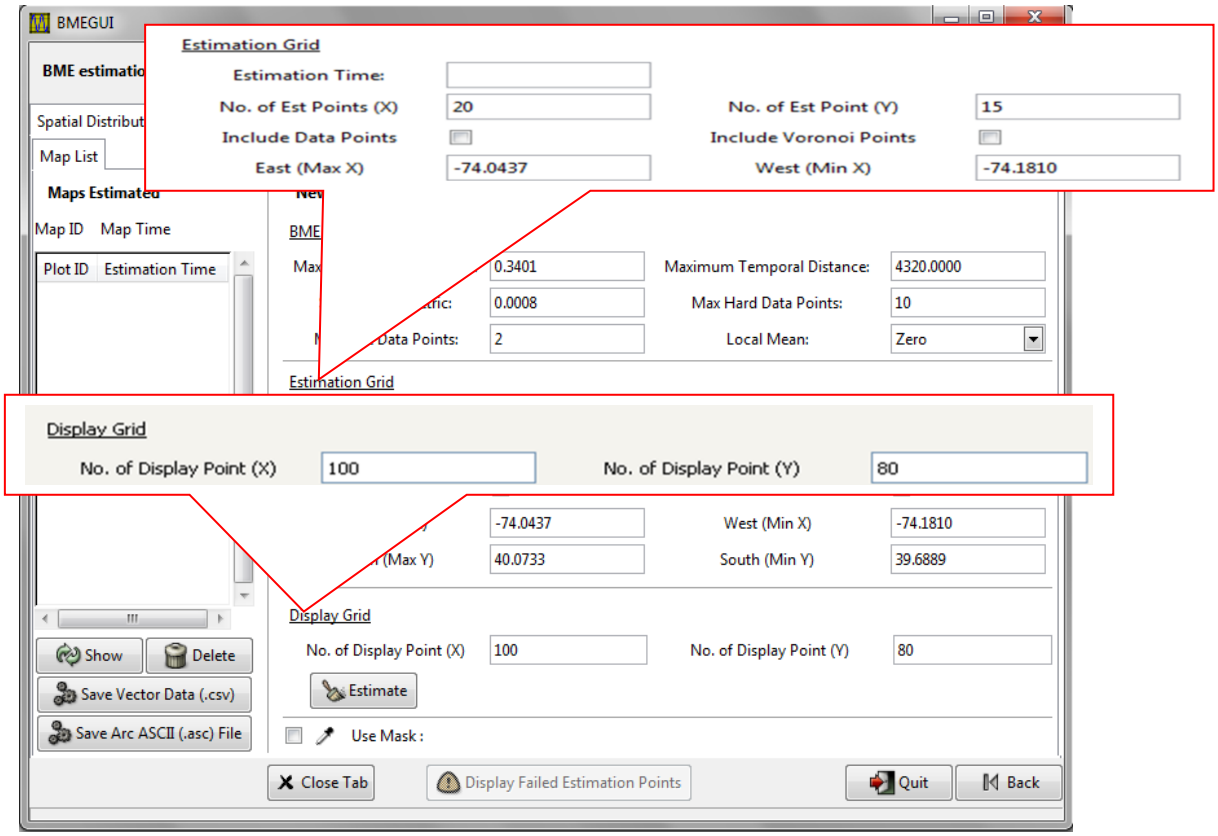


Figure 24: Estimation parameters for the BME spatial estimation

4.8.4 BME Spatial Estimation

As explained in 4.8.2 and 4.8.3, to perform a BME spatial estimation the user needs to specify the BME parameters and the Estimation parameters. Once these parameters are set, the user needs to click on the "Estimate" button on the "Spatial Distribution" tab to create the corresponding map. Then two new tabs are displayed, named "PlotID: xxxx(Mean)" and "PlotID: xxxx(Error)", and a new entry appears on the list in the "Maps Estimated" section (**Error! Reference source not found.**).

User can also provide the mask file of area of estimation to BMEGUI. This is an optional and BMEGUI doesn't consider any mask by default. However, mask file (shape file, with

*.shp extension) should be contained only one single polygon. BMEGUI complains if user provides a mask file containing multiple polygons (poly lines, polygons, poly areas etc).

The map of the BME mean estimated values is plotted on the "PlotID: xxxx(Mean)" tab and the map of BME error variance is plotted on the "PlotID: xxxx(Error)" tab. Maps are displayed by clicking on their corresponding tab (**Error! Reference source not found.**). he list in the "Maps Estimated" section displays all the estimated maps and each entry on the list shows the "Plot ID" and "Estimation Time" of a given map.

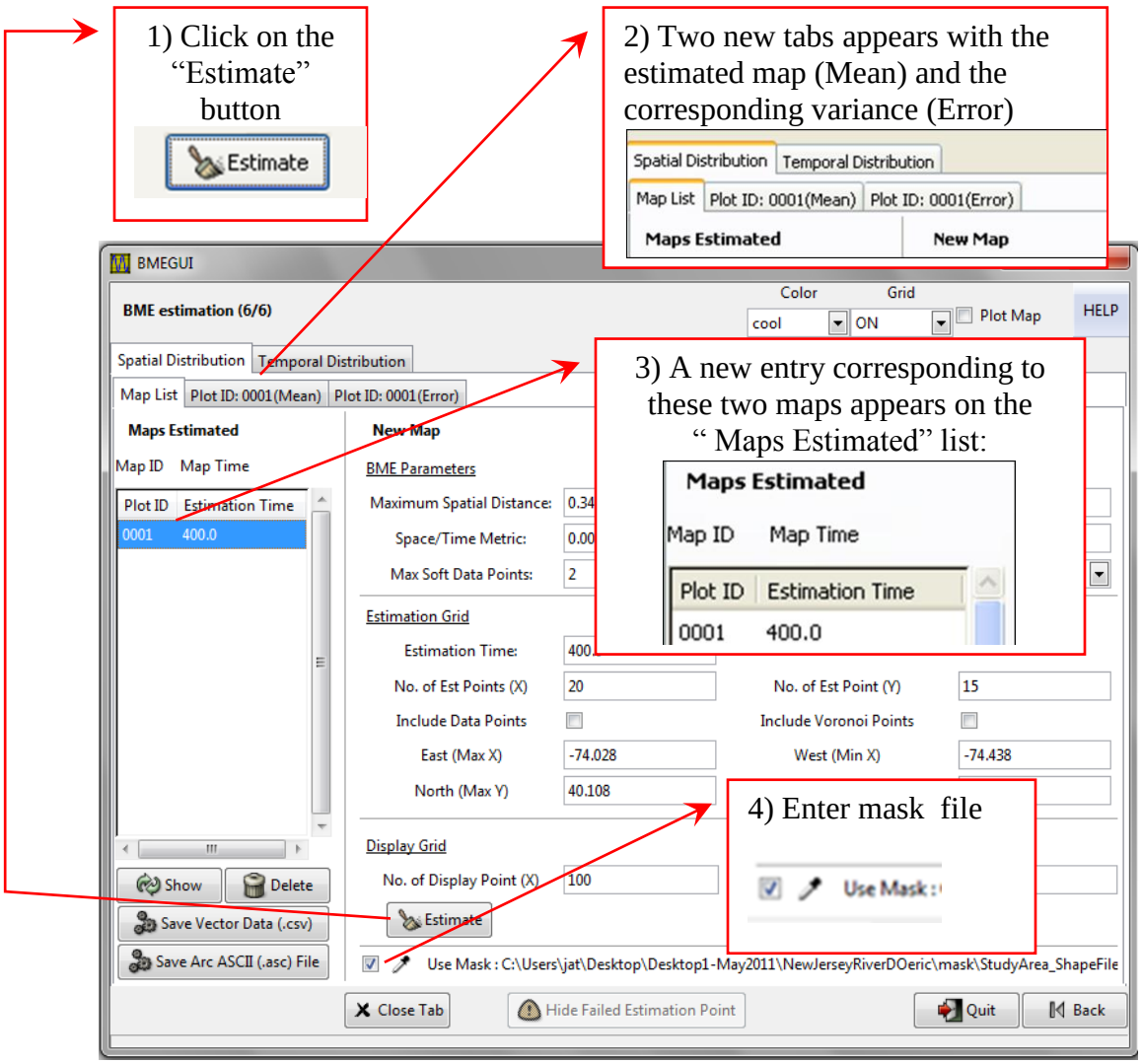


Figure 25: List of BME estimation maps

4.8.5 BME Mask, Map, Grid, and Color Setting

As explained earlier, to perform a BME spatial estimation the user needs to specify the BME parameters and the Estimation parameters. BMEGUI produce BME mean and associated error variance plots and graph. Quality of these graphs and plots can be

improved by selecting appropriate color, and grids from drop down menu. Further user can enhance plots by placing mask and boundary map on the BMEGUI plots by clicking on ‘Mask’ and ‘Map’ buttons.

4.8.6 BMEGUI Home Page and Help

Latest information on BMEGUI updates and new versions can be accessed on BMEGUI web page. Click on ‘HELP’ button seated on top right corner of BME Estimation screen (last screen), and then click on Home page to access BMEGUI web page. You can also contact us if you have any questions, suggestions, and any other BMEGUI related issues. To access the BMEGUI User Manual click on ‘Help’ and then click on ‘User Manual’ button.

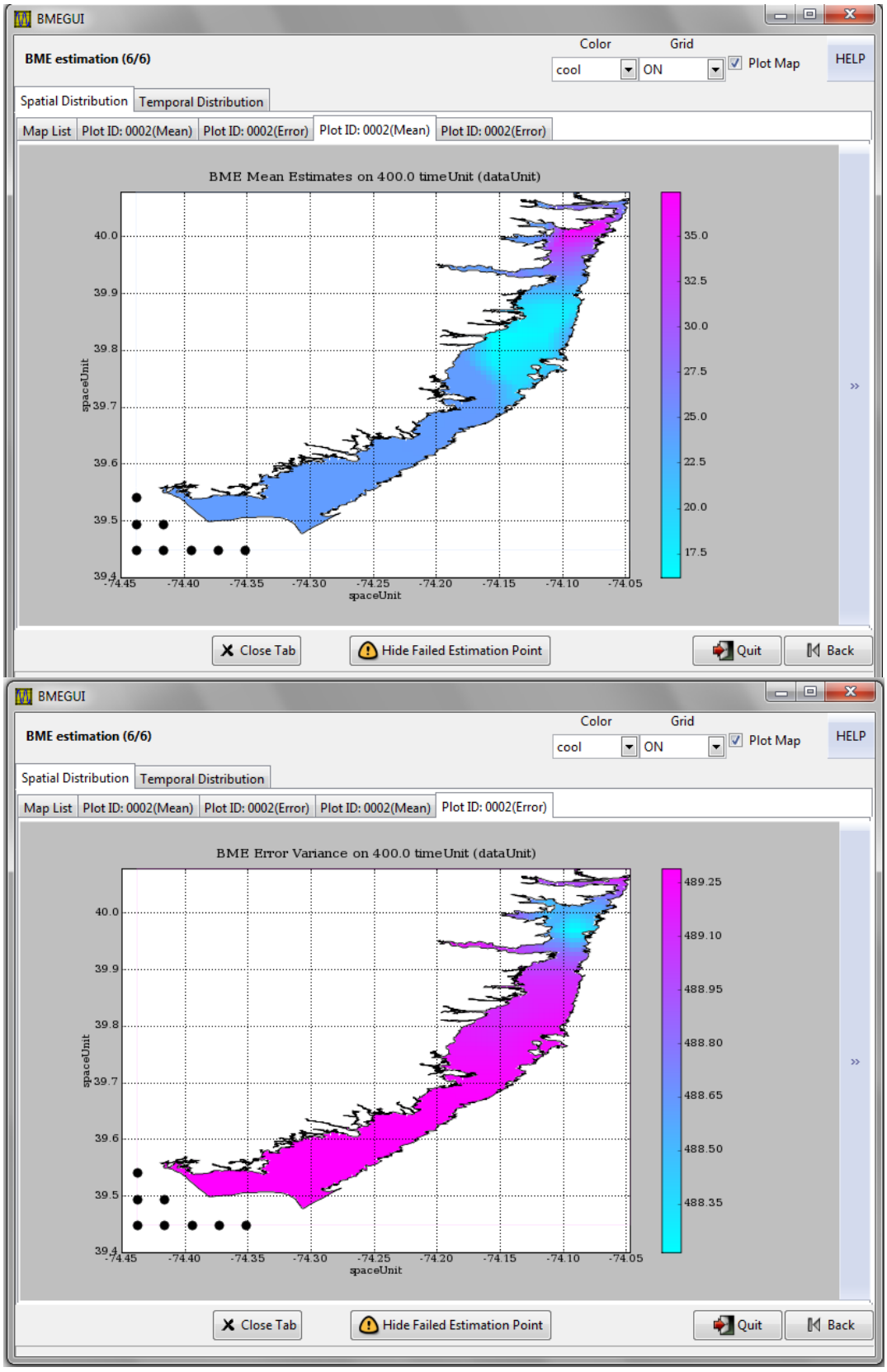


Figure 26: Maps of BME mean estimates and BME error variances

4.8.7 Create Output Files (Point Layer File and ArcASCII File)

As with the spatial distribution plot in Dialog Box 3 (see 4.5.3), the user can also create ArcGIS compactible ArcASCII (*.asc) files from the maps created in Dialog Box 6. The user can also create a vector data file (*.csv), MATLAB and other software compatible, of the BME mean estimate and error variance calculated at each node of the estimation grid.

Arc ASCII format: Arc ASCII file format refers to a specific interchange format developed for ARC/INFO raster's in ASCII format. Arc ACSII format consists of a header that specifies the geographic domain and resolution, followed by the actual grid cell values. Usually the file extension is *.asc.

To create these ArcASCII files for a given map, click on the corresponding entry from the list in the "Maps Estimated" section. Then click on the "Save ArcASCII (.asc) File' File" button or the "Save Vector Data (.csv)" button. Then a message box will appear asking name of the file to be saved as shown (**Error! Reference source not found.**, a), nd place where file has to be saved. After saving file at user specified place, BMEGUI pop ups message saying file ((s) have been created as shown in figure 27 (a) and (b).

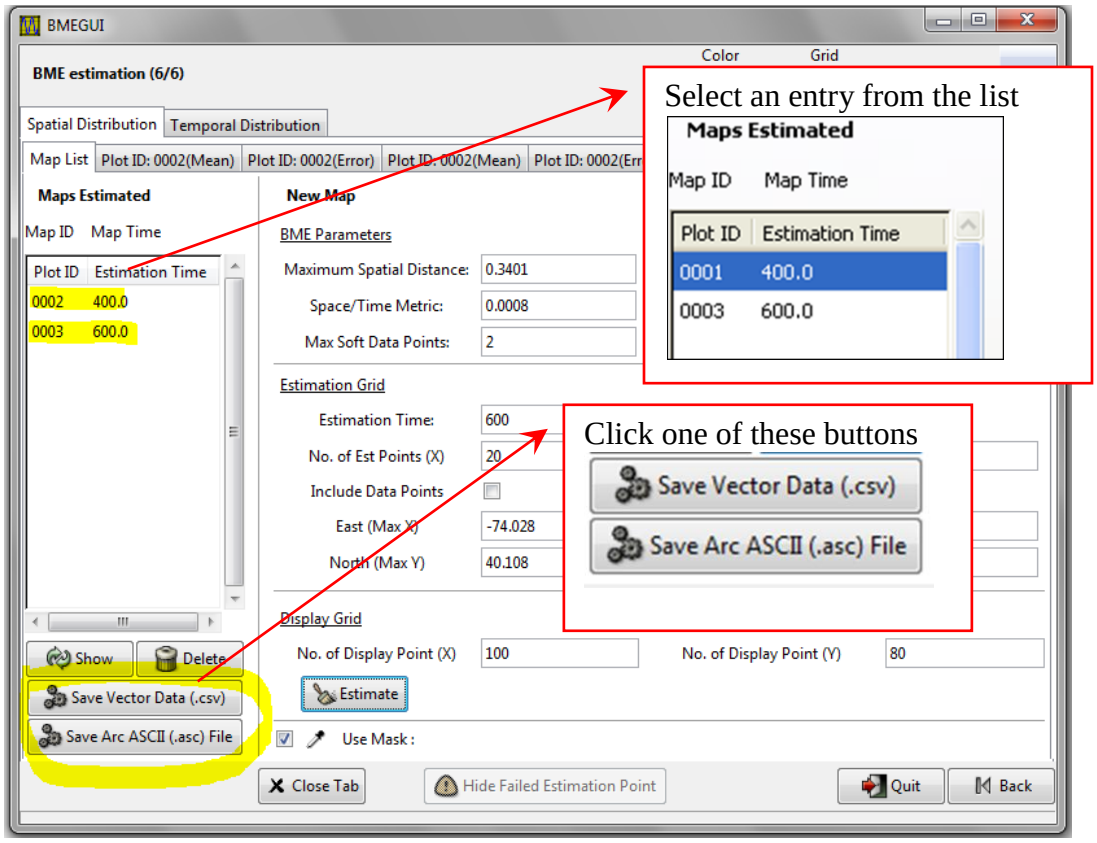
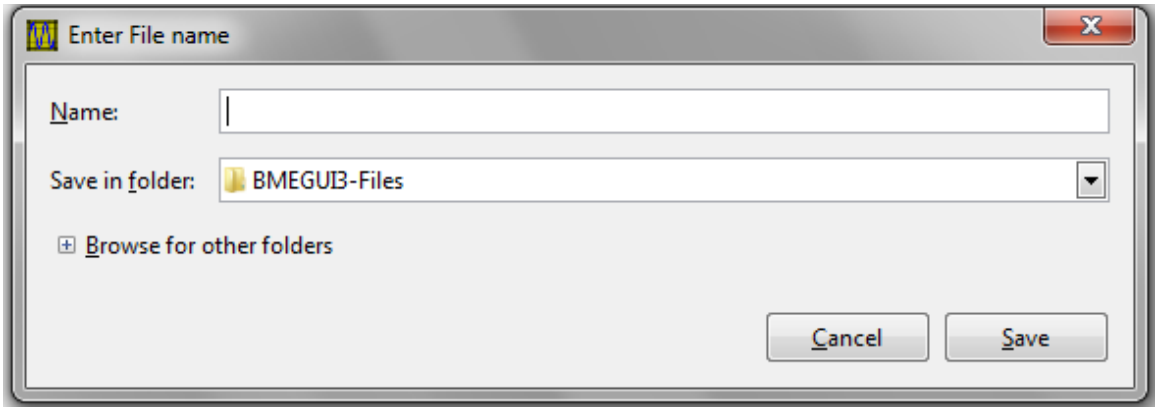


Figure 27: BMEGUI box 6 showing BMEGUI generated maps and buttons for creating ArcGIS compatible and vector data file using BMEGUI output.

(a)



(b) (c)

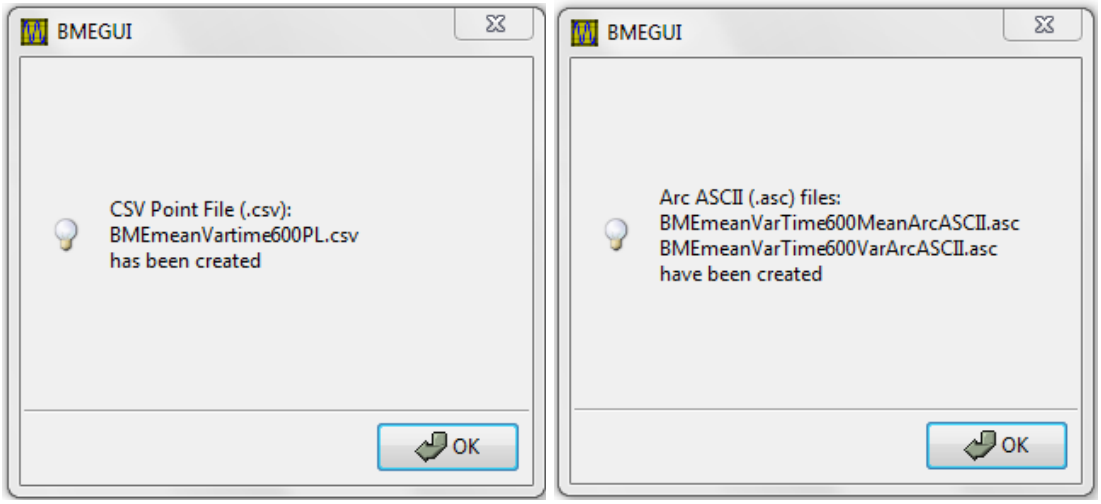


Figure 28: Enter name of the file to save as BMEGUI output (a), message after creating ArcGIS compatible (ArcASCII) file (b), and point layer (csv) files (c)

4.8.8 Estimation Parameters (Temporal Distribution)

In order to obtain the time series plot at specific monitoring locations, the user needs to specify the "Estimation Parameters" and "Display Parameters" (**Error! Reference ource not found.**).

In "Estimation Parameter" section, the user can specify the following parameters.

- ❖ Station ID: ID specifying the monitoring station where the time series needs to be obtained. Select the appropriate station ID from the drop down list.

❖ Estimation Period: User-defined estimation period of the time series.

There is only one parameter in the "Display Parameter" section. This parameter is called the "Scaling Factor", and it is only used for cosmetic effect. This parameter changes the aspect ratio used to display the Gaussian soft data overlaid on the time series plot. The default setting of this parameter is 0.1.

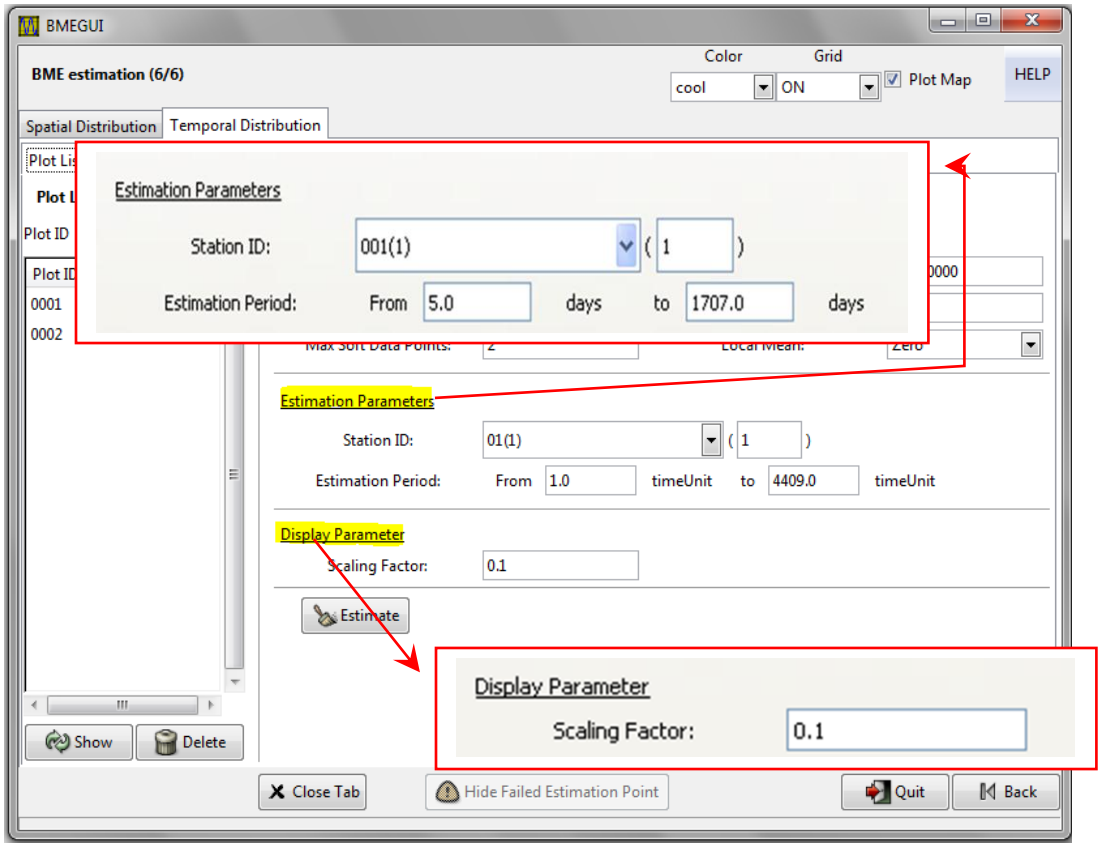


Figure 29: Estimation and Display Parameters used for the BME temporal estimation

4.8.9 BME Temporal Estimation

As explained in 4.8.2 and 4.8.8, the user needs to specify the BME parameters and the Estimation parameters to perform a BME temporal estimation. Once these parameters have been set, the user needs to click on the "Estimate" button on the "Temporal Distribution" tab to perform the estimation. Then a new tab labeled "PlotID: xxxx" is displayed and the corresponding entry appears on the list in the "Plot List" section (**Error! Reference source not found.**).

A plot of the time series is displayed when clicking on the tab (**Error! Reference source ot found.**) corresponding to a specific PlotID. The blue solid line displays the BME mean estimates and the green dotted line shows the lower and upper bounds of the 69% confidence interval (which corresponds to the BME mean estimate ± 1 standard deviation under the assumption of a Gaussian distribution). The blue dots show the hard data, while the red triangles and squares show the hardened soft interval and soft Gaussian data,

respectively. BMEGUI also displays the shape (i.e. either interval or Gaussian) of the PDF describing the soft datum at each soft data point. “Plot List” displays all the estimated time series plots and each entry on the list shows its “Plot ID” and “Station ID”.

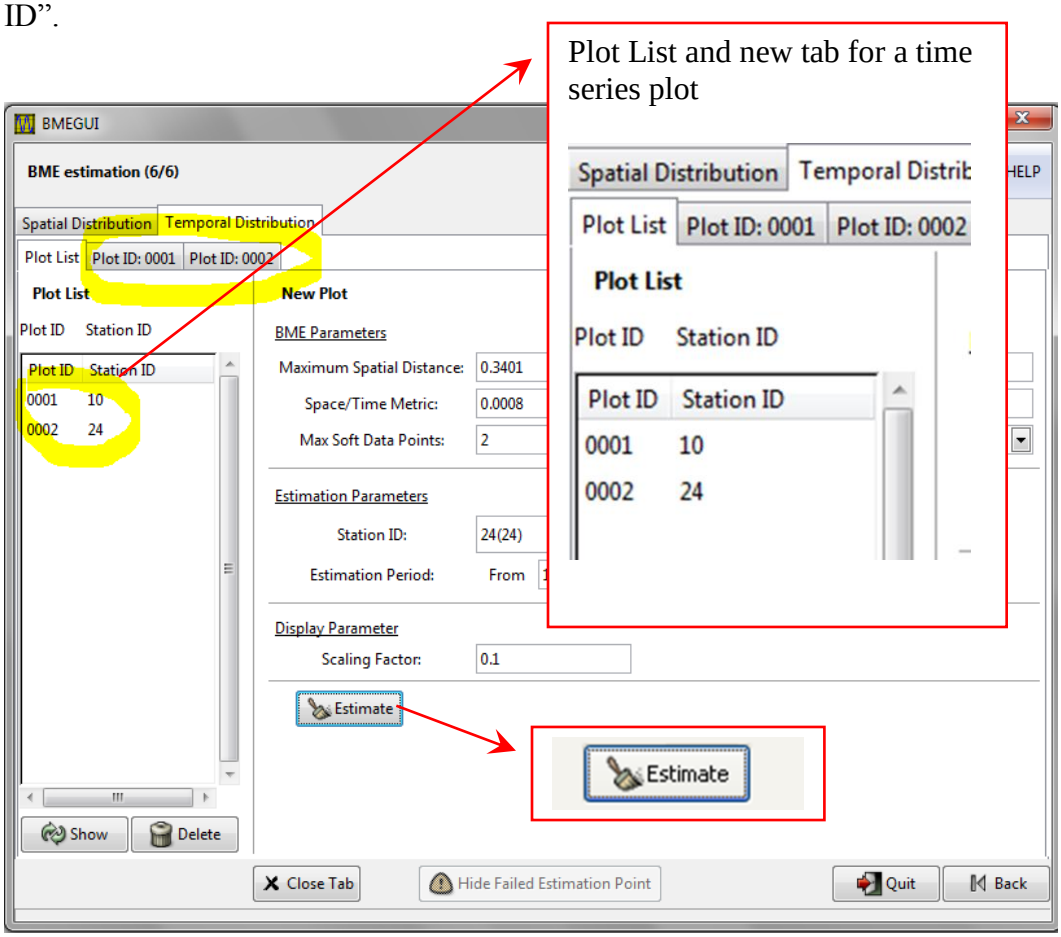


Figure 30: List of estimated time series

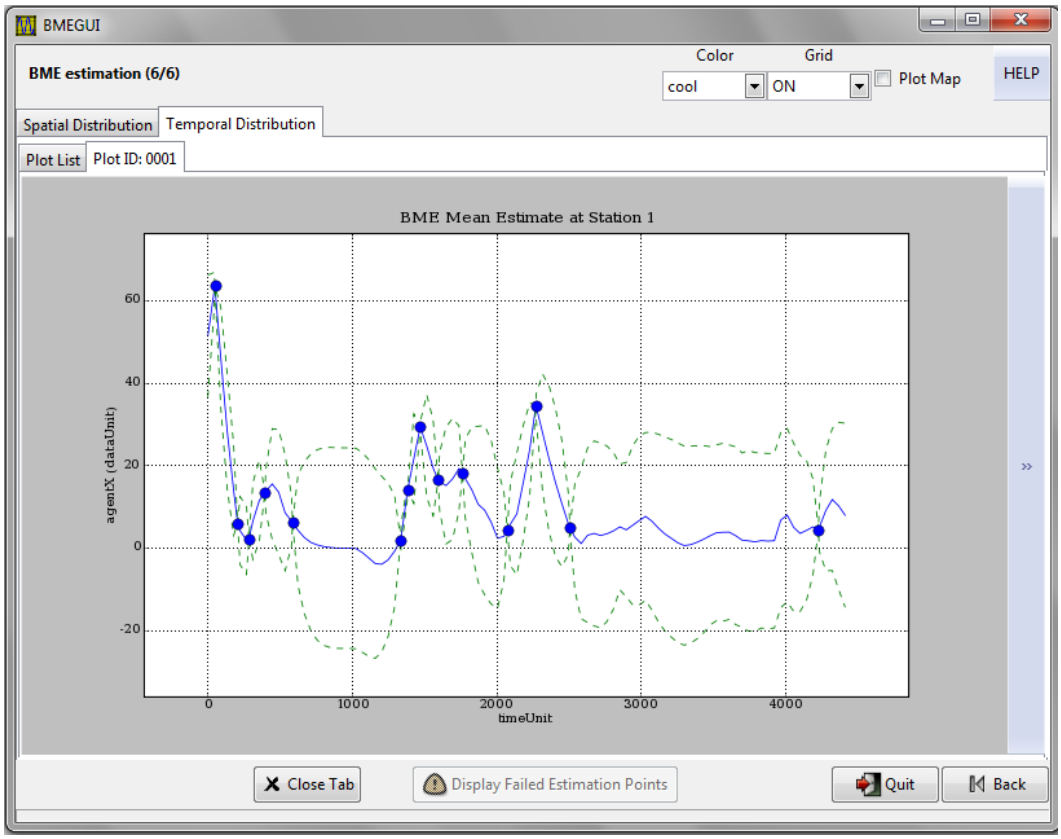


Figure 31: The time series plot at a specific monitoring location

4.8.10 Show, Close, and Delete Maps (or Time Series Plots)

The user can create maps (or time series plots) as many times as s/he wants. Every time the new map (or plot) is created, BMEGUI automatically stores the estimation results. Therefore, the user can temporally close the map (or plot) and redraw the map (or plot) whenever s/he needs it. Moreover, the user can also permanently delete the estimation result (**Error! Reference source not found.**).

To close a map tab (or a plot tab), first click the selected map tab (or plot tab). Then click on the "Close Tab" button and the corresponding tab is hidden. However; the user cannot close the "Map List" tab (or the "Plot List" tab).

To redraw the map (or plot), select the corresponding entry from the map list (or plot list), then click on the "Show" button that is located below the list.

To permanently delete the map (or plot), select the entry from the map list (or plot list), then click on the "Delete" button that is located below the list. A message dialog box will appear, select "OK" to close it.

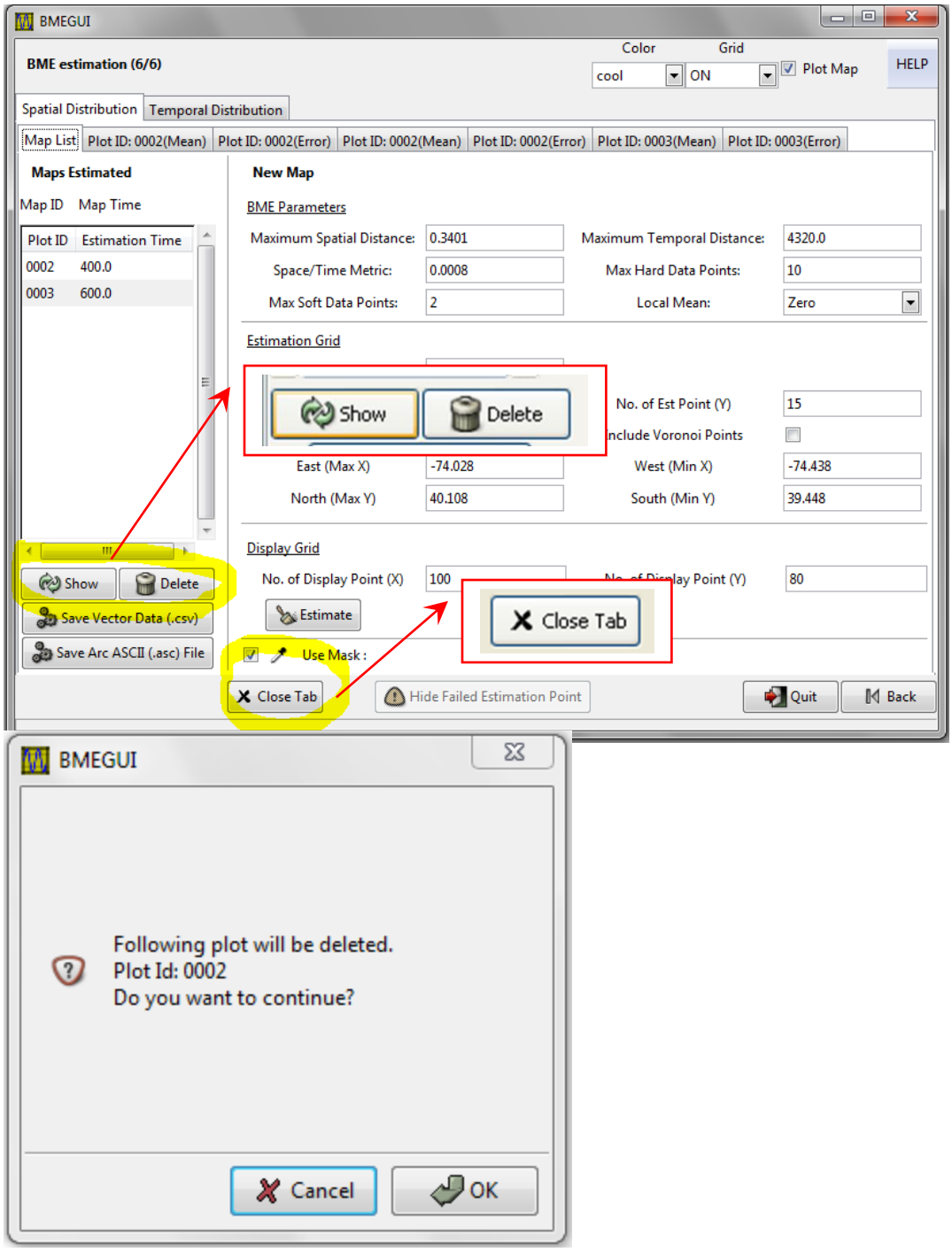


Figure 32: The "Close Tab", "Show", and "Delete" buttons and the message box to confirm the deletion.

4.8.11 Hide and Display Failed Estimation Point

If there are no data in the estimation neighborhood, BME estimation returns a NaN value. The BMEGUI automatically replace the BME estimation result with the average of all the data values. The BME error variance is also replaced with the variance of all the data

values. In order to inform the user about the failed estimation point, the message dialog box telling the number of failed estimation points appears, if there are failed estimation points (**Error! Reference source not found.**). In addition, the failed estimation points re displayed as block dots on the map. The user can hide/display black dots by clicking the "Hide (or Display) Failed Estimation Point" button (**Error! Reference source not ound.**).

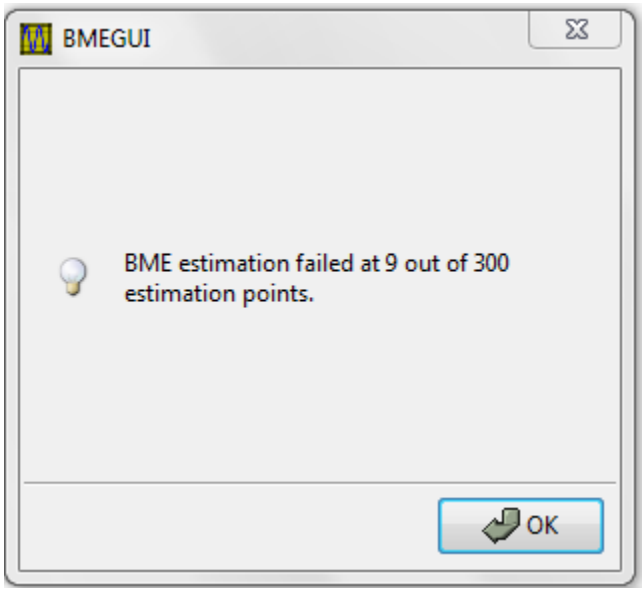


Figure 33: Message showing the number of failed estimation points.

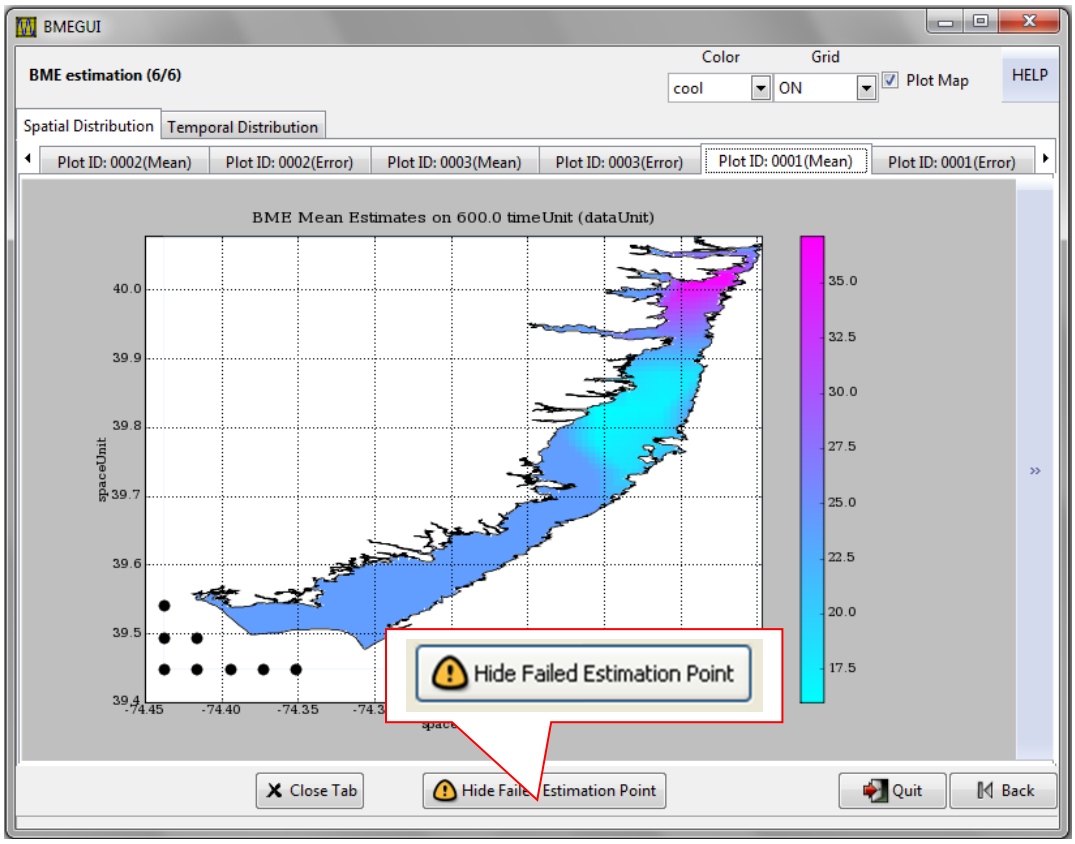


Figure 34: Failed Estimation Points (black dots) on the estimation map

4.9 Quitting from BMEGUI

Each dialog box has a "Quit" button to exit from BMEGUI. When the user presses on the "Quit" button, a message dialog box appears (**Error! Reference source not found.**). ress "OK" to confirm that you really want to quit

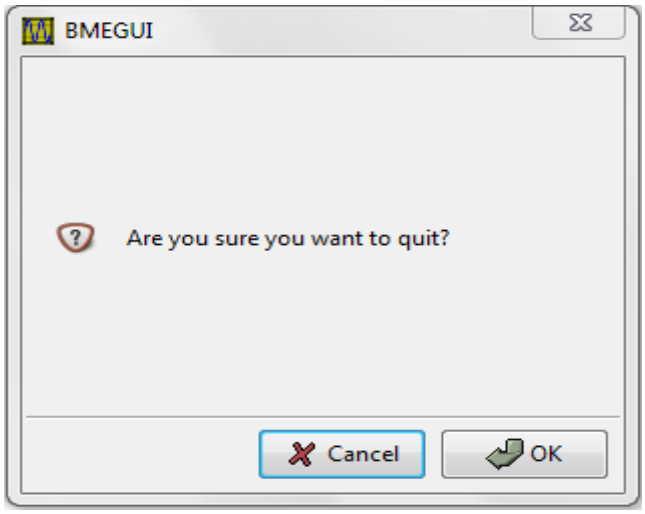


Figure 35: The message dialog box to confirm whether to quit BMEGUI.

5 Interaction with ArcGIS

5.1 Details of ArcGIS Files

As explained in 4.5.3, **Error! Reference source not found.**, and 4.8.7, BMEGUI has unctions to create ArcGIS compatible ArcASCII (*.asc) and CSV (*.cv) files. CSV files can be saved as data base (dbf) file using MS Access to make them ArcGIS compactible. The followings are the list of files that can be created and used with MATLAB and ArcGIS for further analysis.

- Point layer file (*.csv file)
    - Spatial distribution plot
        - Dialog Box 3 (Exploratory Analysis)
        - Data fields: X, Y, T, and Val
    - Spatial raw mean trend (*.csv file)
        - Dialog Box 4 (Mean Trend Analysis)
        - Data fields: X, Y, and Val (Raw mean trend)
    - Spatial smoothed mean trend (*.csv file)
        - Dialog Box 4 (Mean Trend Analysis)
        - Data fields: X, Y, and Val (Smoothed mean trend)
    - BME mean estimate and error variance at estimation grid points
        - Dialog Box 6 (BME Estimation)
        - Data fields: X, Y, Mean (BME mean estimate), and Var (BME error variance)
    - BME mean estimate and error variance at display grid points
        - Dialog Box 6 (BME Estimation)
        - File name: bmeRst(Plot ID).lyr
        - Data fields: X, Y, Mean, and Var
- ArcASCII (*.asc) file
    - BME mean estimate
        - Dialog Box 6 (BME Estimation)
    - BME error variance
        - Dialog Box 6 (BME Estimation)

5.2 Coordinate System of ArcGIS Files

BMEGUI does not define a coordinate system for any of the ArcGIS files created. Therefore, when you add layer file or raster file created by BMEGUI in ArcGIS, the

following warning message will be displayed (**Error! Reference source not found.**). he user can define a spatial coordinate system by using ArcGIS tools.

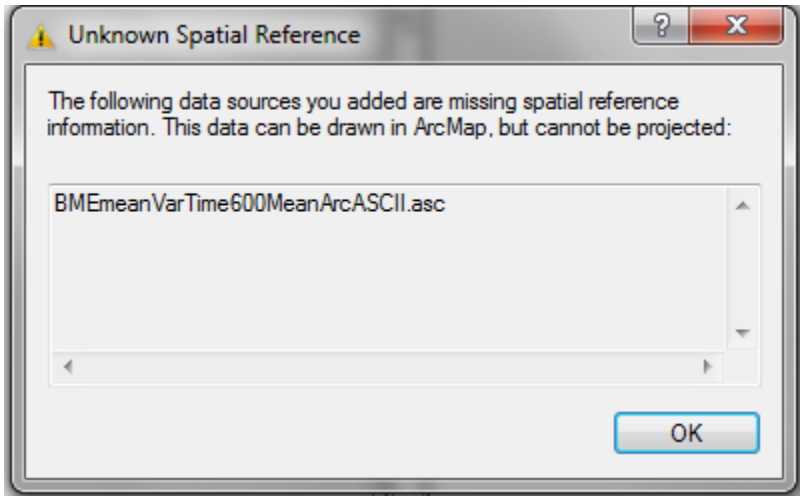

Unknown Spatial Reference
The following data sources you added are missing spatial reference information. This data can be drawn in ArcMap, but cannot be projected:
BMEmeanVarTime600MeanArcASCII.asc
OK

Figure 36: ArcGIS warning message

6 Advanced Topics

6.1 Data Error Handling

BMEGUI can detect and automatically modify the following data errors.

1) The same station ID is assigned to the different geographic locations
2) Different station IDs are assigned to the same geographic location
3) Duplicated measurements

BMEGUI detects and corrects the error in the order listed above. These errors are detected when the user press the “Next” button on Dialog Box 1. BMEGUI displays the message dialog boxes shown in **Error! Reference source not found.** when errors are etected. The user can select whether to accept the BMEGUI default error correction, or to quit the application and correct the error manually. The default error correction methods are listed below.

When BMEGUI detects that the same station ID is assigned to different geographic locations, BMEGUI replaces these different locations with their unique spatial average.

When BMEGUI detects that different station IDs are assigned to the same location, BMEGUI takes the alphanumerically smallest ID as the valid station ID and replaces all the other IDs with it.

When BMEGUI detects duplicated measurements (i.e. measurements made at the same station ID, geographic location and time), BMEGUI takes the average of the duplicated values.

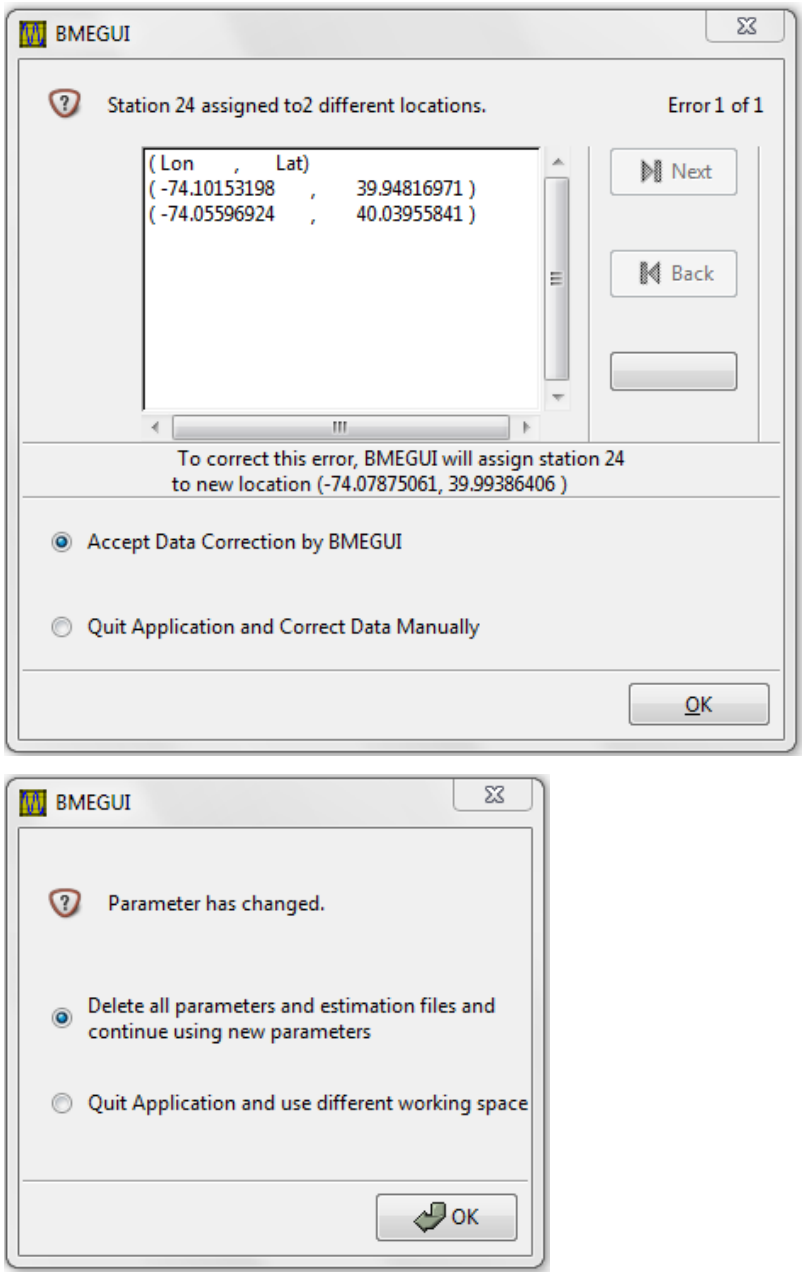

BMEGUI
Station 24 assigned to2 different locations.
Error 1 of 1
(Lon , Lat)
(-74.10153198 , 39.94816971)
(-74.05596924 , 40.03955841)
Next
Back
To correct this error, BMEGUI will assign station 24
to new location (-74.07875061, 39.99386406)
Accept Data Correction by BMEGUI
Quit Application and Correct Data Manually
OK
BMEGUI
Parameter has changed.
Delete all parameters and estimation files and
continue using new parameters
Quit Application and use different working space
OK

Figure 37: The various message dialog boxes that display when data errors are detected.

6.2 BMEGUI Parameter File and Estimation Files

As explained in 3.1.1, when analyzing a specific "Data File", BMEGUI uses the "Workspace" directory to store transient files during BME analysis. The followings are name and description of the parameter file and estimation files that are be automatically created by BMEGUI during the analysis.

- BMEGUI parameter file
  - File Name: (Name of the Data File).ysp

- This file is used to store the all estimation parameters and the intermediary results (including the mean trend and covariance models) generated prior to the BME estimation results produced on “BME Estimation” screen. The information stored in this file is used to reproduce previously obtained intermediary results when the user restarts BMEGUI and specifies the same Workspace and Data File.

❖ BMEGUI spatial estimation files

- File Name: (Name of the data file) + (Plot ID).yme
- This file is used to store the BME spatial estimation parameters and results. Every time the user creates a new estimation map, the PlotID is increased by 1 and a new file is created. These files are used to redraw any map on the map list and restore the corresponding estimation parameters. If the user permanently removes a map from the map list (See 4.8.10), then BMEGUI removes the corresponding file from the workspace.

❖ BMEGUI temporal estimation file

- File Name: (Name of the data file) + (Plot ID).yse
- This file is used to store the BME temporal estimation parameters and results. Every time the user creates a new estimation plot, the PlotID is increased by 1 and a new file is created. These files are used to redraw any plot on the plot list and restore the corresponding estimation parameters. If the user permanently removes a plot from the plot list (See 4.8.10), then BMEGUI removes the corresponding file from the workspace.

❖ Initial parameter files

- File Name: (Name of the data file).py(c)
- This file is used to store initial parameters, such as the number of bins of the histogram, the name of the ArcGIS output files, and other default parameters.

7 Troubleshooting errors

7.1 Data Error file due to an inappropriate new line character

When the data file having an inappropriate "new line" character is specified as the data file in BMEGUI, BMEGUI displays the following error message. (**Error! Reference ource not found.**)

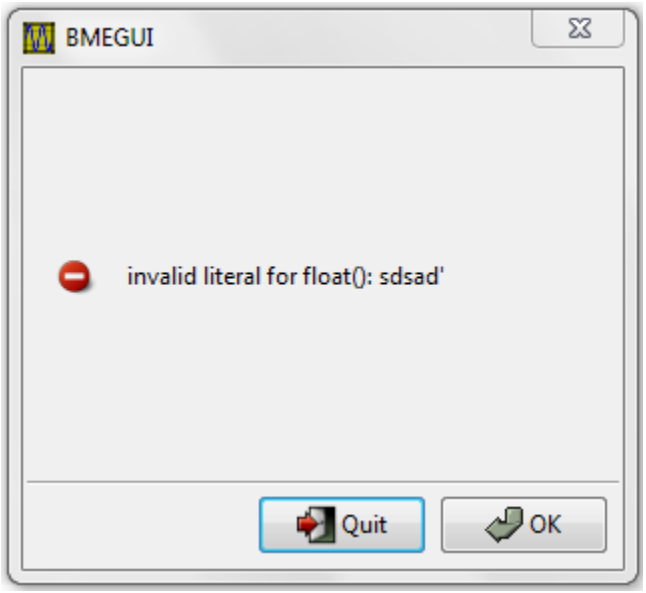


Figure 38: Error message due to an inappropriate new line character

This error might happen when the data file was imported from a Unix or Macintosh machine, or when the data file was created by the "writeGeoEAS" function of BMElib. To fix this problem, use a text editor that is capable of modifying the erroneous "new line" character with the correct "new line" character for Windows. For example you may use the ConTEXT text editor (http://www.contexteditor.org/), as follow

1. Open the data file using context
2. From the "Tools" menu, navigate to "Convert Text To…" and select "DOS (CRLF)" (**Error! Reference source not found.**)
3. Save the file

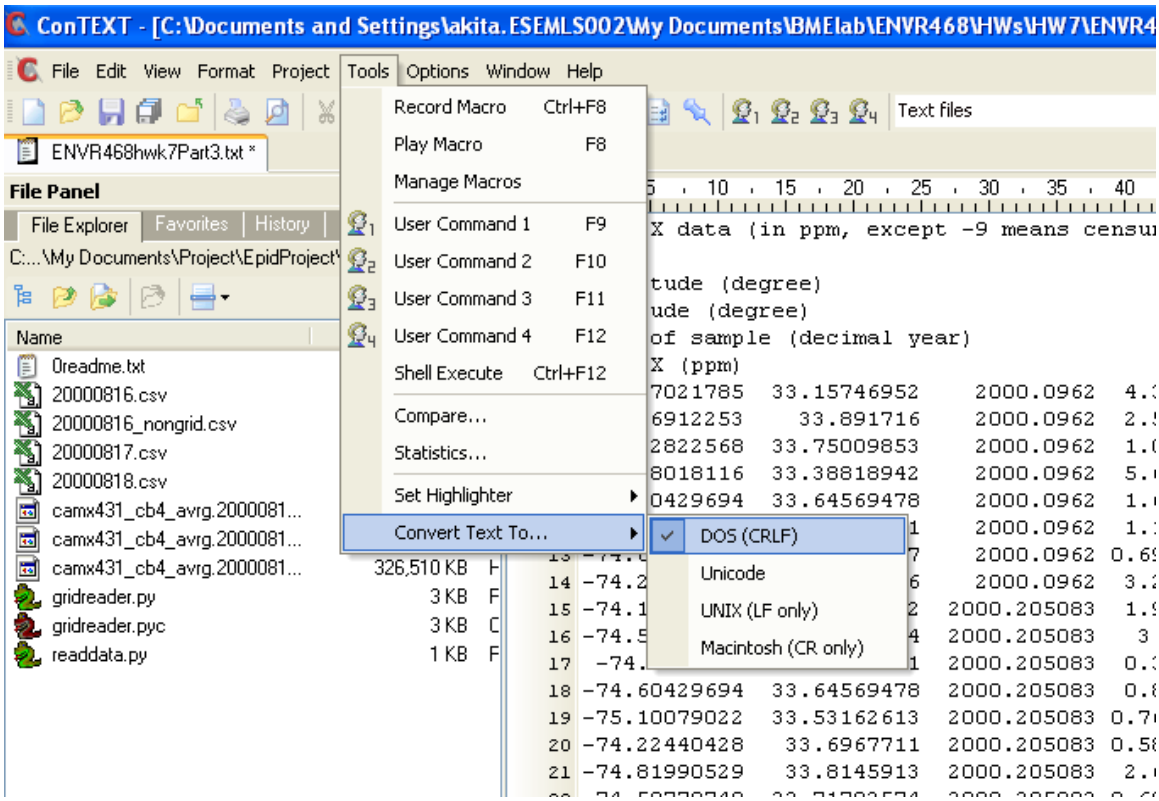


Figure 39: ConTEXT editor